Evanescent Near-Field Optical Lithography: Overcoming the Diffraction Limit

Sharee J. McNab BE *(1st Class Hons.)*

A thesis presented for the degree of
Doctor of Philosophy
in
Electrical and Electronic Engineering
at the
University of Canterbury,
Christchurch, New Zealand.

5 March 2001

ABSTRACT

Concepts of optical resolution limits have been transformed in the past two decades with the development of near-field optical microscopy. Resolutions of $\lambda/40$ have been demonstrated by taking advantage of additional information present in the near field of an object. These resolutions are far higher than what diffraction-limited lens-based optical systems are capable of. Attempts have been made to replicate these resolutions for lithography using a scanning probe based optical equivalent, but these systems suffer from low throughput owing to their serial nature.

A desirable alternative would be replication of all the patterns within a field in a single flood exposure in a manner similar to how optical projection lithography replicates the field of a mask, but with the additional resolution available from working in the near field. This is the basis of evanescent near-field optical lithography, the subject of this thesis. Evanescent near-field optical lithography (ENFOL) brings traditional contact lithography into the near-near field using a combination of conformable masks and ultra-thin photoresists.

This thesis describes a study of ENFOL both experimentally and via electromagnetic simulations to evaluate what the resolution limit might be. The fabrication of membrane masks is described, a key component for the ENFOL exposure. The characteristics of an ENFOL exposure using broad-band light are investigated from exposures into thick resist. These exposures demonstrate the trend of decreasing depth of field as the period of grating structures is reduced.

ENFOL's requirement of a thin imaging photoresist for high resolution lithography complicates the pattern transfer step essential to translate the photoresist image into a useful material for devices. The development of an additive pattern transfer process is described, that utilises a trilayer resist scheme to enable lift-off metallisation. NiCr gratings with periods down to 270nm have been fabricated using this process subsequent to an ENFOL exposure. Wire-grid polarisers consisting of 270nm-period NiCr gratings on glass substrates have been fabricated and their polarisation properties measured at visible wavelengths.

Simulation results of exposures of sub-wavelength grating structures are presented that investigate the fundamental limit to resolution for contact lithography techniques such as ENFOL. A full-vector, rigorous electromagnetic simulation technique, the mul-

tiple multipole program is used to provide information about the near field of sub-wavelength gratings. The potential for $\lambda/20$ resolution is indicated; a tantalising prospect for optical lithography and well below the diffraction limit of conventional optical projection-based lithographies. Perhaps the most critical parameter for an evanescent exposure, the depth of field, was characterised and a linear relationship shown between the depth of field and grating period. The effect of parameters such as grating duty cycle, absorber material and thickness on the exposure are observed with the intention to optimise the experimental setup.

Interesting interference phenomena are observed in simulation results for exposures where the effective exposure wavelength is equivalent to the grating period. In particular a period halving occurs in the transverse magnetic polarisation due to interference of the first diffracted orders. A novel interference technique — evanescent interference lithography is proposed that takes advantage of an enhanced period halving at an exposure wavelength corresponding to a grating resonance.

PREFACE

This dissertation describes research undertaken in the Department of Electrical and Electronic Engineering at the University of Canterbury between February 1997 and February 2001. I am grateful to my supervisor for suggesting the topic of research, his enthusiastic supervision and his efforts in creating a microfabrication facility and research group virtually from scratch. When I first began my doctorate, the laboratory was equipped with an optical microscope, photoresist spinner and mask aligner. My initial experimental work was some microwave analogy work inspired by a shortage of equipment for microfabrication. Today while our facilities are still not vast, a respectable collection of equipment has been acquired for pursuing world-class research.

Being involved at the inception of the Nanostructure, Engineering, Science and Technology group had its advantages and disadvantages. I gained useful experience in setting up processes for using the equipment and hopefully gained a good understanding of the tools. However it was also a time consuming procedure with few direct benefits in terms of original research; for while the roads being trodden were not new, an infrastructure and history of processes which accumulates over time had not been established.

The "black art" of fabrication, while more under my control after some years "at the bench", has lost none of its mystery for me. My initial expectations of processes working first time and my impatience that they did not has been tempered somewhat with experience. The frustrations however, make the achievements more satisfying. As a lover of variety I have enjoyed the multifaceted nature of this topic, covering aspects of microwave, microfabrication and electromagnetic simulation even though to the untrained eye it may look as though I was just drawing lines.

Aspects of the work described in this dissertation have been published as follows:

R. J. Blaikie and S. J. McNab. Evanescent interferometric lithography. *Applied Optics*, 40(10):1692–1698, April 2001.

R. J. Blaikie, M. M. Alkaisi, and S. J. McNab. Evanescent near field optical lithography for low-cost nanofabrication. In *Proceedings of the Advanced Research Workshop on Semiconductor Nanostructures*, pages 47–49, Queenstown, New Zealand, 5-9 February, 2001.

S. J. McNab, R. J. Blaikie, and M. M. Alkaisi. Analytic study of gratings patterned by evanescent near field optical lithography. *J. Vac. Sci. Technol. B*, 18(6):2900–2904, 2000.

M. M. Alkaisi, R. J. Blaikie, and S. J. McNab. Low temperature nanoimprint lithography using silicon nitride molds. In *Micro- and Nano-Engineering 2000*, Jena, Germany, 18-21 September, 2000.

S. J. McNab and R. J. Blaikie. Exposure characteristics of gratings for evanescent near-field optical lithography. In *Progress in Electromagnetics Research Symposium Proceedings*, page 801, Cambridge, Massachusetts, USA, 5-14 July, 2000. The Electromagnetics Academy.

S. J. McNab and R. J. Blaikie. Contrast in the evanescent near field of $\lambda/20$ period gratings for photolithography. *Applied Optics*, 39:20 – 25, January 2000.

M. M. Alkaisi, R. J. Blaikie, S. J. McNab, R. Cheung, and D.R.S Cumming. Sub-diffraction-limited patterning using evanescent near-field optical lithography. *Appl. Phys. Lett.*, 75(22):3560–3562, 1999.

M. M. Alkaisi, R. J. Blaikie, and S. J. McNab. 70 nm features on 140 nm period using evanescent near field optical lithography. *Microelectronic Engineering*, 53:237–240, 2000.

S. J. McNab and R. J. Blaikie. Investigating the fundamental limit to resolution in the evanescent near field. In *Proceedings of Image & Vision Computing New Zealand*, pages 97 – 102, Christchurch, NZ, 30-31 August, 1999.

M. M. Alkaisi, R. J. Blaikie, and S. J. McNab. Nanolithography using wet etched silicon nitride phase masks. *J. Vac. Sci. Technol. B*, 16(6):3929–3933, 1998.

S.J. McNab, R.J. Blaikie, M.M. Alkaisi, D.R.S. Cumming, and R. Cheung. Extending the resolution of optical lithography using the evanescent near field. In *Proceedings of the Fifth Electronics New Zealand Conference*, Dunedin, 31 August - 1 September, 1998.

ACKNOWLEDGEMENTS

First and foremost, I wish to express my sincere gratitude to my supervisor Dr Richard Blaikie, for his support and guidance throughout this thesis and for suggesting the topic of research. I am also appreciative of his continuing optimism and his ability to always highlight the positive. I would also like to thank my co-supervisors, Dr David Cumming and Dr Rebecca Cheung. I was sorry to see you both leave but glad of your input while you were here, David for introducing me to electron beam lithography and Rebecca for introducing me to reactive ion etching. I am grateful to Dr Maan Alkaisi for sharing with me his experience in experimental research, particularly in the early stages of this work, and for many useful discussions. Thanks to Dr Roger Reeves for introducing me to spectroscopy in the Department of Physics. I would like to thank Helen Devereux for all her valuable help in the microelectronics laboratory; Dermott Sallis, Steve Downing and Nick Smith, for help in resuscitating equipment at times, and Pete Lambert for his machining skills.

I would like to acknowledge the financial support from the Canterbury Doctoral Scholarship and a scholarship from the New Zealand Federation of University Women, as well as assistance from the Royal Society of New Zealand, and the IEEE in the form of travel grants.

A special thanks must go to my fellow students both in the Nanostructure Engineering Science and Technology group, and to my postgraduate colleagues for providing necessary distractions from the task in hand, particularly Cressida Harding, Katharine Holdsworth, Timo Bretschneider, Tim Drysdale, Bifeng Rong and Valerie Leung. I'd also like to specially thank my cousin Kathleen, also completing her PhD, for being a great listener and a great talker. There are many friends who I would like to acknowledge, for joining me in escapades and making my thesis years enjoyable, especially Larisa Clarke, Bente Clausen, Richard Cranstone, Lucette Dijkstra, Jon Hunt, Maria Moran, Vanessa Sherlock and Fiona Young.

I would like to express my deepest gratitude to Alan for his love and patience, for proof-reading much of this thesis, and for being my prime accomplice in adventure. Finally, I'd like to thank my parents for many years of love and support and for instilling in me the confidence to try anything, and the fortitude to see it through.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

AFM	atomic force microscope
BRR	background resist reduction
DFB	distributed feedback laser
DI	de-ionised water
DOF	depth of field
EAM	embedded-amplitude mask
EBL	electron beam lithography
EIL	evanescent interference lithography
EM	electromagnetic
ENFOL	evanescent near-field optical lithography
FDTD	finite difference time domain
FEM	finite element method
GMT	generalised multipole technique
HMW	high molecular weight
IC	integrated circuit
IPA	isopropyl alcohol
JPEG	Joint Photographic Experts Group
LCM	light coupling mask
LER	line edge roughness
LMW	low molecular weight
MIBK	methyl-isobuty-ketone
PMMA	poly(methyl methacrylate)
PR	photoresist
RCWA	rigorous coupled wave analysis
RIE	reactive ion etching
RPM	revolutions per minute
SEM	scanning electron microscope
SNOM	scanning near field optical microscope
STM	scanning tunnelling microscope
TE	transverse electric
TEM	transverse electric & magnetic

TM	transverse magnetic
UV	ultra-violet

Chapter 1

INTRODUCTION

"Good things come in small packages" is an age-old adage that the microelectronics industry has taken to heart. Lithography, the technique that defines the circuit detail, is largely responsible for the ever decreasing "package" size. This thesis explores an optical lithography variant that addresses the quest for smallness in a very low-cost manner.

Microelectronic circuits have become pervasive in all aspects of modern life, from communications to transport and consumer goods. The microelectronics industry came about as a result of two major developments. Firstly, the invention of the transistor by Bardeen, Brattain and Shockley in 1947. Secondly, a process was developed to integrate transistors onto a single silicon chip — the integrated circuit (IC), to create a manufacturable technology. The development of the IC was made in the late 1950's independently by Kilby from Texas Instruments and Noyce then of Fairchild Semiconductor [1]. Kilby's pioneering invention of the IC was recognised in 2000 by the award of the Nobel Prize in Physics, acknowledging its significance in catalysing the electronics era.

Since the development of the IC there has been staggering progress in miniaturisation of the constituent transistors. Many benefits are associated with miniaturisation, including the ability to fabricate more complex circuits, which, coupled with improvements in device speed has meant that larger numerical problems can be tackled. The reduced weight and volume of circuits has also enabled their use in applications such as air and space travel. But perhaps the biggest benefit was the IC's huge impact on the cost of electronics manufacture. An initial prediction in 1965 by Dr Gordon Moore that the number of devices in an IC would double each year became an impressive reality and was known thereafter as Moore's Law [1]. In 1975 he revised his prediction to a more sedate doubling every 18 months to two years [2]. The Semiconductor Industry Association adopted Moore's Law as a roadmap to aid planning and co-ordination of the inter-disciplinary effort to achieve each new technology node. This is now an international effort and Moore's law has become a somewhat self-fulfilling prophecy. Figure 1.1 plots the minimum linewidth achieved versus year, showing the logarithmic

decrease in linewidth. The increased device density of chips is illustrated in terms of the number of memory bits on dynamic random access memory (DRAM) chips from 16 kb in 1977 to a predicted 4 Gb in 2003 [3].

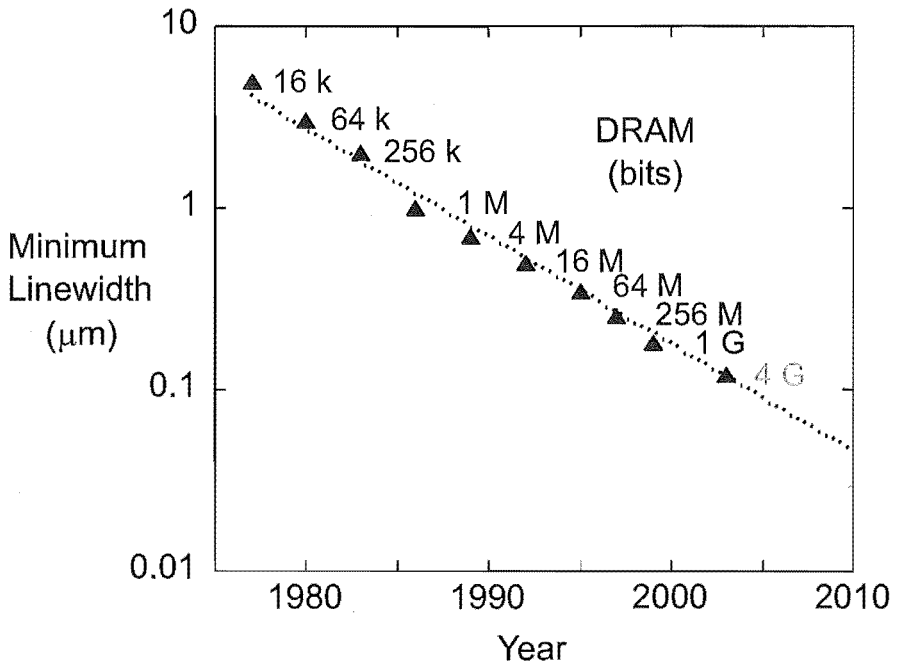


Figure 1.1 Illustration of Moore's law showing the logarithmic decrease in minimum feature size and the concomitant increase in density for DRAM.

The two dimensional process used to fabricate IC's is known as the planar fabrication process. It enables many devices to be engineered onto a single piece of silicon by a sequence of processing steps that include lithography, growth/deposition, etching, and doping. These are illustrated in Fig. 1.2. With this seemingly simple tool set, complex circuits of more than twenty layers can be built, with transistors numbering in the millions on a single chip.

Lithography is a key enabling technology in IC manufacture; it is responsible for defining the patterns which form transistor gates, metal interconnects, contacts, and vias. The speed, performance and economics of ICs is inextricably linked to the minimum lithographically printable feature size, and over the past three decades intense effort has been focused on its reduction. Over half the increase in clock speed of each new generation of microprocessor is attributable to process improvements [5]. As the device density is increased, the processing cost per device is improved. The lithography system costs are a significant portion of IC cost, typically one third of the cost to process a wafer to completion [6].

Optical projection lithography is the mainstay of the semiconductor industry today. Optical wavelengths in the UV and deep UV are used to expose a light sensitive chemical

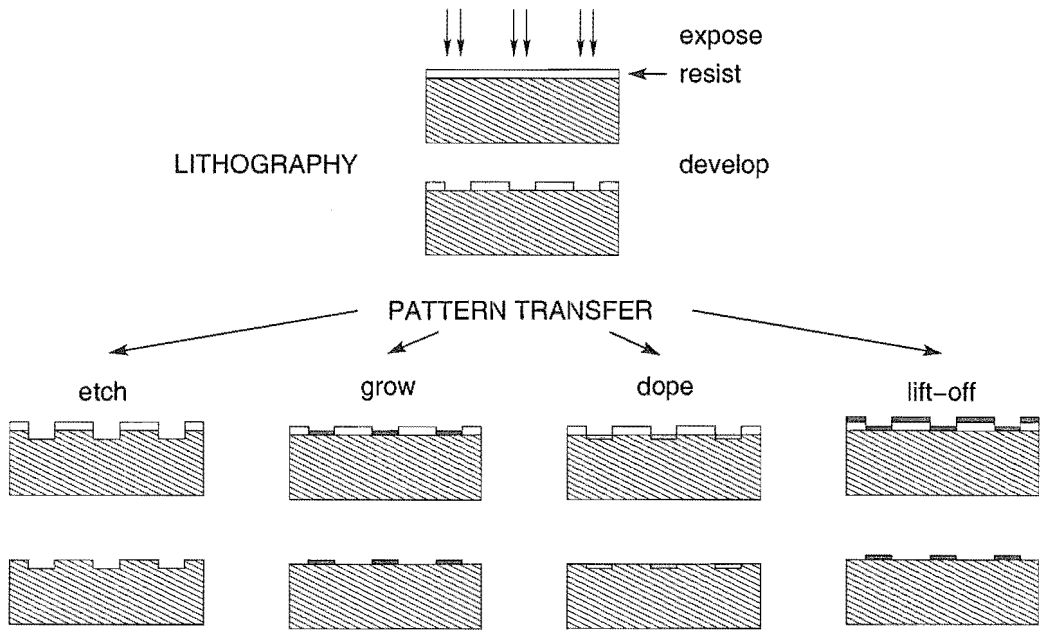


Figure 1.2 The planar fabrication processes, lithography followed by pattern transfer [4].

film (a photoresist). The light passes through an amplitude mask and is demagnified and focused through a system of lenses. This technique is discussed in more detail in Section 1.1 as well as how the reduction in critical dimensions has been achieved.

The current era of optical lithography is forecast to end around 2005, either preceeding or following the introduction of the 100nm technology node [3]. Next generation lithographies (NGLs) such as extreme ultra-violet lithography (EUVL), electron projection lithography, x-ray and ion projection lithographies are waiting in the wings, vying to be the replacement for optical systems. Currently EUVL [7] and electron projection lithography schemes such as scattering with angular limitation projection electron-beam lithography (SCALPEL) [8] and projection reduction exposure with variable axis immersion lenses (PREVAIL) [9] appear to be the favourites, however the outcome is by no means certain.

While the semiconductor industry is the biggest user of lithography for making logic circuits, random access memory (RAM) and the like, there are a wealth of alternative applications for which lithography also plays an essential role. For example lithography is used for making magnetic memory, diffractive optical devices such as lenses and gratings, and for making holograms in the security industry. Lithography is also becoming increasingly prevalent in biotechnology areas for making sensors, or cell sorting and trapping devices for cell analysis [10, 11]. With all these different applications, each possessing different requirements and tolerances, it is likely that no single lithography solution will be appropriate for all. Niche markets may turn to alternative

lithographies to those used by high volume IC manufacturers, to fulfil their needs. The following sections describe optical lithography, the current technique used for IC fabrication and briefly discuss the alternatives for the next generation lithography before describing some low-cost lithography alternatives.

1.1 OPTICAL LITHOGRAPHY

Optical lithography uses light having a wavelength λ, anywhere from the visible range to the deep ultraviolet ($\sim$700 – 150 nm) to expose a light sensitive resist usually through an amplitude mask. It is generally considered to be a diffraction limited technique because as the mask feature sizes gets smaller, diffracted light degrades the resultant image, limiting the resolution to somewhere in the order of the exposing wavelength.

Optical projection lithography continues to be the dominant lithography used by the semiconductor industry today. This is despite experts' predictions as early as the 1970s that optical lithography would not be able to keep pace with the resolutions demanded of it. Optical lithography has had many guises since its inception, originally starting out as a contact printing technique. Yield problems due to contamination arising from hard contact of the mask and substrate encouraged a move from contact to proximity printing where a small gap was introduced between the mask and substrate. Since the 1980's projection based systems have gained popularity. This thesis introduces another variant of optical lithography, evanescent near field optical lithography (ENFOL). It is closest in form to contact printing techniques but with some important improvements.

1.1.1 Contact Printing

Contact printing was the mainstay of semiconductor lithography prior to projection lithography and is still often used in many research and production processes. Typically a broad-band UV exposure source is used to expose a photoresist coated wafer, through a 1× mask, typically chrome on glass. The mask and wafer are generally held in hard contact using a vacuum. The set-up for contact printing is illustrated in Fig. 1.3(a). Good fidelity with high throughput at low cost has been demonstrated with this technique. The industry moved away from contact printing owing to the following drawbacks. Particle contamination was transferred between the mask and substrate causing low device yield and damaging the mask; this in turn lowered device yield further with successive exposures. A lack of flatness of the masks and substrates as well as particle contamination can result in variable contact causing variations in the linewidth. Feature misalignment is also a problem owing to in-plane distortions.

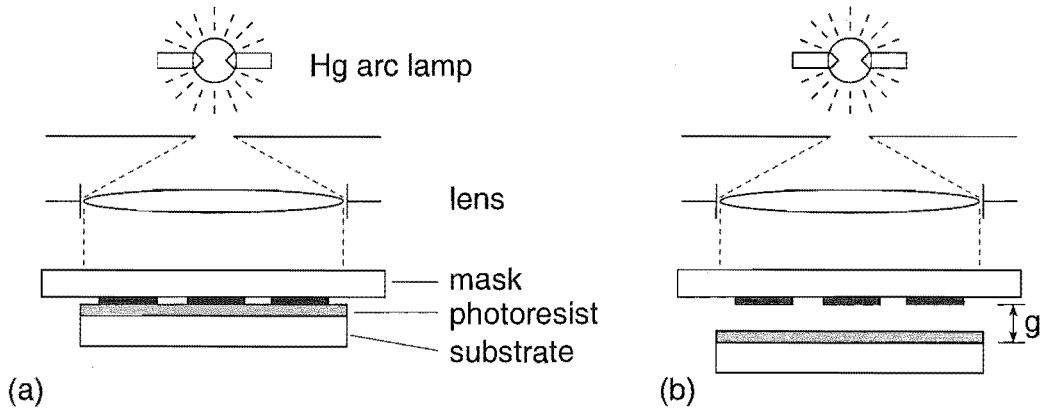


Figure 1.3 Schematics of optical printing techniques (a) contact printing and (b) proximity printing [12].

1.1.2 Proximity Printing

Proximity printing was an attempt to overcome the mask contamination issues prevalent in contact printing. A small gap was introduced between the mask and substrate, typically 5 to 50 μm, as illustrated in Fig. 1.3(b). Various methods of setting the gap are used, physical spacers on the periphery of the mask, or in some mask aligners a hard contact is initiated then a stepper motor used to back the wafer off a set distance. The inability to control wafer and mask flatness results in some variability in the gap between different regions that is translated to variations in the printed linewidths.

Increasing the gap has a detrimental effect on resolution, owing to diffractive broadening. The resolution l_{min} (in terms of the minimum linewidth) of contact/proximity printing schemes is given as [12]

$$l_{min} \cong \sqrt{\lambda g} \tag{1.1}$$

where λ is the illumination wavelength, g the gap between the mask and substrate and including the photoresist thickness. For example with a 10 μm gap and $\lambda = 400$ nm, l_{min} in the region of 3 μm can be expected. If a gap is necessary, then better resolution can be achieved by placing a lens in the gap; this is the basis of optical projection lithography.

1.1.3 Optical Projection Lithography

In the 1980s, the microelectronics industry adopted optical projection lithography. Huge advances in optics technology enabled the generation of a uniform image over large field sizes. Monochromatic sources were introduced to allow precise tuning of lenses for one wavelength and to avoid chromatic aberrations. A simplified optical

projection scheme showing the main components is illustrated in Fig. 1.4.

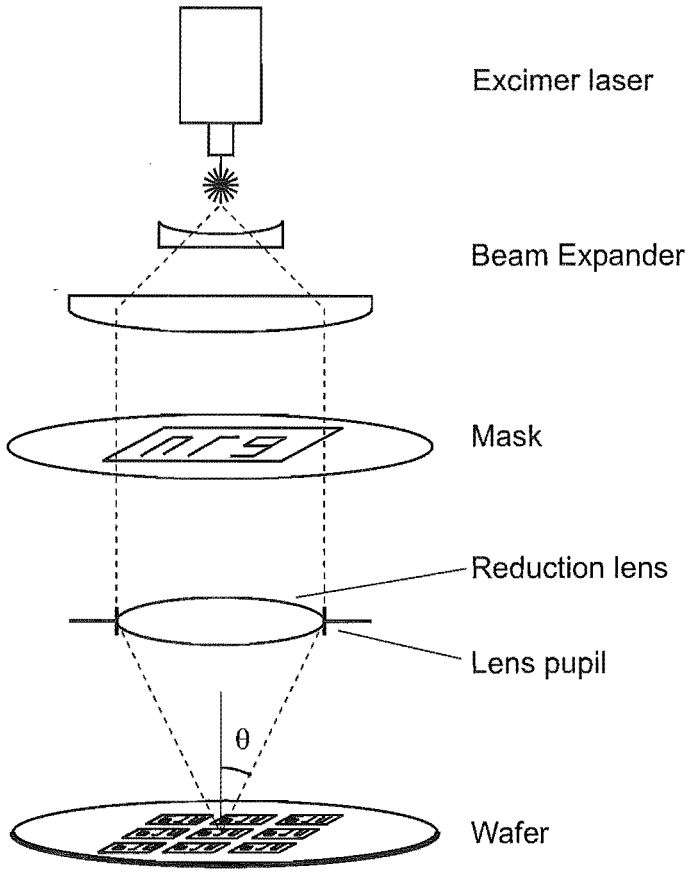


Figure 1.4 Schematic of a simplified optical projection system.

The resolution l_{min} capability of projection systems is well understood and defined as [12]

$$l_{min} = \frac{k_1 \lambda}{\mathrm{NA}} \tag{1.2}$$

where λ is the source wavelength and NA the numerical aperture of the optics system. The NA defines the ability of the optical lenses to collect the diffracted light and is given by $\mathrm{NA} = n \sin\theta$, where n is the refractive index of the surrounding material, and θ half the angle subtended by the lens aperture. A proportionality constant k_1 encompasses the effect of the resist system and resolution enhancements. Theoretically $k_1 \geq 0.25$ [13], but typically values are much greater than this. To improve the resolution two main avenues exist, decreasing the wavelength or increasing the NA. To date the most significant advances have been made by reducing the source wavelength. Initially the G-line of a mercury arc lamp (436 nm) was employed, the wavelength was then reduced to the H-line (405 nm) followed by the I-line (365 nm). To provide adequate high-brightness sources in the deep UV, new excimer lasers were developed, a KrF laser at

248 nm and an ArF laser at 193 nm. An F_2 laser at 157 nm is also being developed. However, while reducing the wavelength improves resolution, other aspects become more difficult. The silica lenses start to become increasingly absorbing in the deep UV and this absorption leads to compaction which results in abberations and focus changes. Materials such as CaF_2 have lower absorptions but they are thermally sensitive and hygroscopic [14]. The resists too become increasingly absorbing, signalling a trend towards thinner resists. Less conventional top surface imaging resists or multilayer resist systems may be required if satisfactory single layer resist materials cannot be found. The depth of focus (DOF) reduces with wavelength as [12]

$$\mathrm{DOF} = \frac{\lambda}{2\,\mathrm{NA}^2}. \tag{1.3}$$

Again, this requires thinner resists and places stricter constraints on the surface planarity, particularly for multi-layer patterning.

1.2 NEXT-GENERATION LITHOGRAPHIES

Optical projection lithography is predicted to finally run out of steam somewhere around 2008, at the 70 nm technology node. This is when the next-generation lithography (NGL) is scheduled to take over for semiconductor fabrication [3]. Projection electron-beam lithography and extreme ultraviolet lithography are the current favourites to be the NGL, after recommendation by International SEMATECH in 1999 for continued development funding. Serious issues are yet to be resolved however, so the outcome is by no means certain. In this section some of the key NGL contenders are outlined.

1.2.1 X-Ray

X-ray lithography uses x-rays as the radiation source with wavelengths typically between 0.4 and 4.5 nm. The most common x-ray lithography technique is known as proximity x-ray, which employs a shadow mask similar in principle to that of proximity/contact printing [15,16]. As x-rays are particularly energetic, the absorber must be made of a thick high atomic mass element such as Au or W to provide adequate shadowing. A low atomic mass material is required for the mask substrate to obtain high transmission. Membrane masks made of materials such as Si, SiC or SiN have been employed [16], similar to the membrane masks described in this thesis for evanescent near-field optical lithography. High intensity x-rays can be generated by synchrotron sources. These sources are expensive, but this is consistent with the trend of increasing capital costs for all the NGL's. Proximity x-ray lithography provides a parallel exposure with the possibility of high throughput. Resolutions of 30 nm have been reported using proximity x-ray lithography [16]. As x-rays have a short wavelength compared

to optical wavelengths, the effects of diffraction are reduced. As with optical proximity techniques, the resolution is a function of the mask-to-sample gap and can be described by Equ. 1.1.

At one stage x-ray lithography appeared to be poised to replace optical lithography, however, this status has since changed as other technologies such as electron projection lithography and EUVL have gained favour. Uncertainties due to mask making are largely responsible for the decline in popularity of x-ray lithography. The combination of thin membrane masks and high-aspect-ratio absorbers required for high resolution are difficult to manufacture and can cause patterning errors due to in-plane stresses. A 1× mask is also not popular as it places tighter constraints on mask making and inspection. Additional lithographic considerations include exposure blur due to photo-electrons emitted during x-ray absorption and penumbral blur due from exposure from point-like sources.

1.2.2 Direct Write Electron Beam

Direct write electron beam lithography (DWEBL) was born from developments in scanning electron beam microscopy and cathode-ray tube technologies, and the discovery of an electron sensitive resist [17]. Direct write electron beam lithography delivers a focused beam of electrons directly to a resist coated substrate in an evacuated chamber. The desired pattern is defined via a vector scan, using a beam blanking mechanism to blank the beam between pattern writes. Electrons have very small effective wavelengths, typically in the order of $0.04 - 0.004\,\text{nm}$ at the voltages of interest. Resolution is limited not by the effective electron wavelength however, but by a combination of the beam diameter and proximity effects. The beam diameter is determined by the quality of the electron optics for focusing the beam and more fundamentally by space-charge effects caused by the mutual repulsion of the electrons as they are forced closer together. The proximity effect describes the unwanted secondary exposure that occurs due to scattering in the resist and backscattered electrons from the substrate [18]. This results in linewidth broadening and exposures that are pattern dependent.

DWEBL currently plays a pivotal role in making photo-masks for optical lithography and is an important research tool. It is a high resolution, flexible tool but its serial nature means that it has an inherently low throughput, making it unsuitable for high volume microelectronics manufacture. A number of attempts have been made to overcome the throughput problem. Variable-shaped beam writing was developed to effectively print a number of pixels in parallel [19], but this method still has low throughput compared to optical steppers.

An alternative solution is to write patterns in parallel with a number of electron emitters [20]. Etec Systems is developing such a system, electron-beam microcolumn lithography, that uses arrays of miniature electron-beam columns (micro-columns) that

operate at 1-2 keV [21]. Each microcolumn consists of a Schottky emitter with its own set of miniature electrostatic lenses and blanker. Proof-of-concept operation has been demonstrated using 16 columns with a $2 \times 2\,\mathrm{cm}^2$ footprint. An advantage of the low voltage design is a reduction in proximity effects. However in order for the low energy electrons to expose the resist, thin resist layers are required. A suitable resist for low voltage exposure and subsequent pattern transfer is still being sought.

1.2.3 Projection Electron-Beam Lithography

Projection electron-beam lithography is an attempt to provide a parallel electron exposure. An incoherent parallel beam is used to illuminate a scattering mask with a field size in the order of $1 \times 1\,\mathrm{mm}^2$ at the mask. The field size is kept small to control optical aberrations, and a step-and-scan like writing strategy is employed to expose different regions of a mask. Electron optics are used to produce a demagnified image of a mask region at the wafer. Two systems are currently under development, scattering with angular limitation projection electron-beam lithography (SCALPEL) [8, 22] and projection reduction exposure with variable axis immersion lenses (PREVAIL) [9]. The SCALPEL mask consists of a very thin silicon nitride membrane with low-atomic-number material and strongly scattering high-atomic-number material areas. The low-atomic-number material weakly scatters the electrons, while those passing through the high-atomic-number material are scattered at higher angles. The mask is separated into numerous windows with supporting struts in between for thermal and mechanical support. The size of the window is the field size of the electron exposure. The mask and wafer are then stepped to expose all portions of the die [23].

1.2.4 Extreme Ultraviolet Lithography

Extreme ultraviolet lithography (EUVL) uses wavelengths between $10 - 14\,\mathrm{nm}$ as the exposure source in a projection-based system. It is similar in principle to optical projection lithography with major differences owing to the strongly absorbing behaviour of materials in this range of the electromagnetic spectrum. As a consequence, the optical lenses must be designed from reflective rather than refractive elements. To achieve high reflectivities, elements are made up of multilayer thin-film coatings, typically alternating layers of Mo and Si, to give resonantly enhanced reflection at the design wavelength and angle of incidence. Reflective masks as opposed to transmission masks must also be employed; these consist of patterned absorbers on multilayer reflectors. Resists also have high absorbance at EUV wavelengths so thin imaging resist layers are required. Resists such as silylated single layer resists and multilayer resists are under investigation [24].

Major challenges facing EUVL include the high degree of perfection required of the mirrors that comprise the projection system, in terms of surface figure (shape)

and surface finish. Surface roughness less than 0.25 nm is required [24]. Design of the multilayer reflectors is complicated by the large angular range of incident light. Variations in the angle of the incident light result in variations in the reflectance which in turn affect the resultant image. Mask defect repair is another challenge to be met before a manufacturable solution is reached. Suitable resist systems are still being sought with high contrast and sensitivities compatible with high throughput at EUV wavelengths. Additionally, the resists must have suitable properties for subsequent pattern transfer, such as good etch resistance.

1.3 REVIEW OF LOW COST ALTERNATIVE LITHOGRAPHIES

The battle of the giants over which technology will claim supremacy as *the* NGL is typified by capitally intensive solutions with massive research efforts being directed to develop them. Meanwhile, at the other end of the spectrum, an array of low cost, high resolution lithography alternatives are being researched such as nanoimprint, micro-contact printing, conformable contact photolithography and lithography using binary phase shift masks. While many of these lithographies may not be appropriate for the high volume, high yield requirements of the semiconductor industry, there are a host of alternative applications for which the high capital costs of optical projection lithography or NGL lithographies are prohibitive. Examples of these applications include flexible displays, diffractive optics, security holograms, or printing on curved surfaces. Matching some of these low cost alternative lithographies with suitable niche applications is an exciting prospect.

Some of the main low cost lithography techniques are reviewed below. They can be differentiated into non-optical contact and optical contact methods. There is some degree of commonality amongst the techniques in that they all require physical contact of a mask and substrate, all with 1× masks, apart from the case of the binary phase shift mask. They all have the advantage of being parallel replication processes which means that throughput is not a problem, however there are still many challenges to be overcome with regards to alignment and the minimisation of distortions.

1.3.1 Nanoimprint Lithography

Nanoimprint creates a thickness contrast in a resist layer by physically pressing a mask with a surface relief pattern into the resist. The principle is illustrated in Fig. 1.5. Initially nanoimprint was proposed as a low cost solution for creating patterned magnetic nanostructures for increasing magnetic storage densities. While imprint technology has been around for several decades (this is how compact disks are made), structures of sub-micron dimensions had not been imprinted. In 1995 Chou *et al.* demonstrated sub-25 nm patterning using this technique [25] and since then has re-

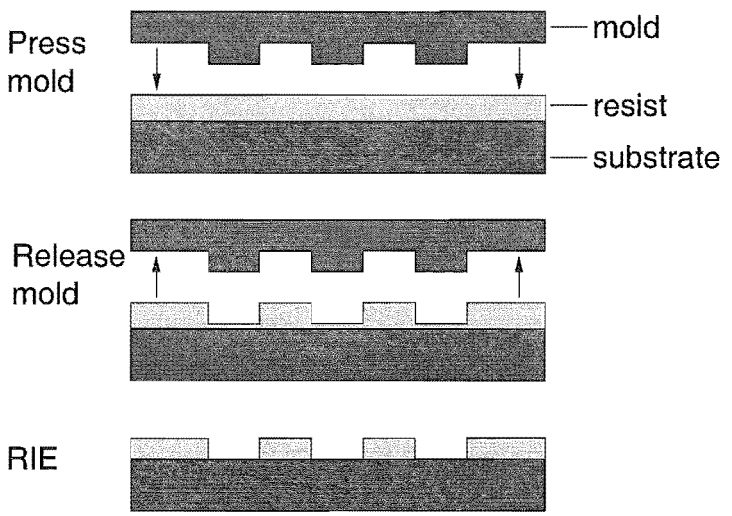


Figure 1.5 Schematic of the nanoimprint lithography process. Imprint a mask with surface relief to create a thickness contrast in a resist. Remove the mold, then RIE to transfer the pattern to remove residual resist in compressed areas.

ported 10 nm dots on a 40 nm period [26, 27]. Imprint molds are made from a silicon wafer, with a surface-relief pattern defined in a SiO_2 layer. The substrate, a silicon wafer, is coated in PMMA and the two heated to well above the glass transition temperature of the PMMA (105°C) so that the polymer becomes a viscous liquid and able to flow. The mold and substrate are pressed together at pressures of about 1900 psi and held together until the temperature drops below the glass transition of the PMMA before being separated. Surface sticking problems have been reported [28], with many groups using mold release agents to prevent the PMMA sticking to the mold [28, 29]. An alternative solution that the author has been involved with is the use of Si_xN_y as the mold material. Si_xN_y has demonstrated nanoimprint without sticking problems [30], additionally it is well known for being a hard and durable material, which may improve problems with the structural integrity of sub-10 nm mold structures with high aspect ratios. Low temperature nanoimprint has been reported [30, 31]. Imprint at low temperatures may be advantageous for alignment (an issue yet to be addressed) as the pattern displacements that can occur between the mold and substrate owing to disparities in their thermal coefficients are reduced. Low temperatures also mean that the imprint is compatible with processes having a low thermal budget.

1.3.2 Microcontact Printing

Microcontact printing is closely related to traditional printing. A mask or stamp is literally 'inked' then contacted to the substrate. Typically the stamp is made of an elastomeric material such as poly(dimethyl siloxane) (PDMS) which is cured on a pat-

terned master to create the surface relief (similar to a nanoimprint mask). Elastomeric stamps have the desirable properties of being conformal and able to accommodate rough surfaces, and are chemically inert.

Self-assembled mono-layers (SAMs) have been used as the ink since they form a contiguously ordered layer on the stamp. Microcontact printing has been demonstrated with SAMs of alkanethiolates on gold [32–34] and silver [35] surfaces. Despite the thinness of the alkanethiolates, high wet etch resistance has been demonstrated to allow pattern transfer, for example a SAM of hexadecanethiol protects the gold it coats during a cyanide wet etch [32–34]. Nanometre scale patterning in the order of 80 nm has been demonstrated, using a octadecyltrichlorosilane SAM which was printed onto an amorphous silicon layer which was then etched in KOH with the SAM as the resist film [36].

Microcontact printing has the advantage of not being limited by diffraction or depth-of-field constraints. Printing on curved surfaces has been demonstrated with microcontact printing; an amplitude photomask was printed on an optical fibre surface [37]. The topography of the stamp is vital; steep high-aspect ratio structures are desirable to obtain sharp boundaries between inked and non-inked regions. As the feature size is reduced, the structural integrity on the stamp can be affected. An aspect ratio of 0.3 (feature depth compared to width) is suggested as a minimum for transferral of sub-100 nm patterns [33].

Some interesting work has been reported using these elastomeric materials as a mold for molding organic polymers instead of stamping. An organic polymer such as polyurethane is brought into contact with the patterned mold master under mechanical pressure to transfer the pattern [35]. In essence it is the same as nano-imprint except that the mold is deformed during the molding process. For example a lateral compression might be applied to reduce the feature sizes in the organic polymer, or forces applied to obtain non-planar patterning.

There are some notable deficiencies of the elastomeric masks in the areas of registration for accurate pattern placement and distortions due to the elastic deformation of the masks [38]. These deficiencies are equally applicable to other techniques utilising the elastomeric masks, including microcontact printing, molding, photolithography with binary phase shifting masks and photolithography with light coupling masks.

1.3.3 Binary Phase-Shift Masks

Binary phase-shift masks can be used to pattern sparse nano-scale features without the necessity of a high resolution mask. The mask is transparent to light and intensity contrast is created at phase boundaries. Light arriving from different sides of the phase boundary arrives out of phase, interfering destructively to create a region of low intensity. The principle is illustrated in Fig. 1.6. To achieve maximum contrast, the

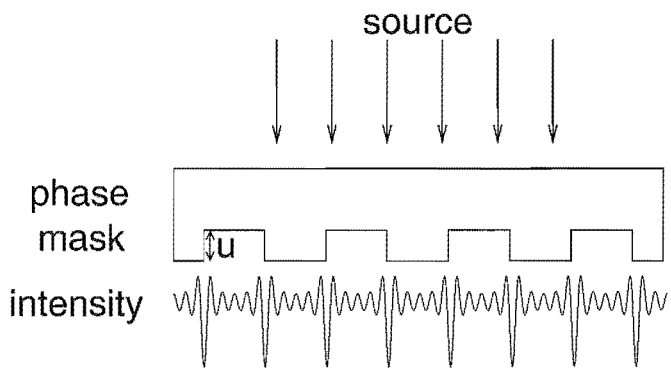


Figure 1.6 Schematic showing the intensity resulting from an exposure with a binary phase shifting mask.

phase boundary which is created by a surface relief in the mask is chosen to be a depth equating to a phase change of an odd multiple of π. The depth d of the surface relief to give a π phase change is given by

$$d = \frac{\lambda}{2\triangle n} \tag{1.4}$$

where λ is the illumination wavelength, and $\triangle n$ the difference in refractive index between the mask material and the surrounding medium.

Sub-100 nm resolution has been demonstrated using this technique [37, 39]. Rogers *et al.* used PDMS elastomeric phase masks, made by casting and curing PDMS onto a silicon master with topographical patterns defined lithographically in photoresist [37]. The mask was then exposed with UV light in conformal contact with a photoresist coated substrate. Alkaisi *et al.* used silicon nitride masks etched in HF to obtain an isotropic profile in the surface relief pattern [39]. The silicon nitride membranes were also dry-etched with CHF_3 and Ar to obtain vertical profiles [40]. The flexible membrane masks were held in intimate contact with a 500 nm thick photoresist coated substrate. The width of the resultant patterns is not controllable although it is repeatable. Simulations performed by the author suggest that by controlling the etch depth with the isotropic profile some degree of control over the pattern width may be possible [39]. Further details of this work are reported in Appendix A. Experiments have also demonstrated that the distance between the mask and substrate significantly affects the pattern width [40].

The relationship between the mask and resultant image is not a simple one-to-one relationship as the previously described low-cost techniques have been. The intensity image is akin to the inverse of the differential of the phase mask with respect to height; an area of low exposure is evident at points corresponding to an abrupt π phase change.

The remarkable ability to produce nanometre-scale features with large mask fea-

tures that do not require high resolution patterning is the main advantage of this technique. The main disadvantage is the limited pattern geometries that can be created; only closed geometries can be patterned easily. For example, an arbitrary line-section is impossible to create in a binary phase mask as the phase boundary cannot end abruptly, but must meet itself in a closed geometry. A tapering of the phase boundary is a solution, but then the mask profile is complicated from a simple binary phase to include graded profiles.

The binary phase shifting mask should not be confused with the related alternating phase-shift masks first proposed by Levenson [41]. Alternating phase-shift masks consist of a standard amplitude mask with a transparent phase-shifting material added to alternating apertures of the mask. The thickness is chosen to create a π phase change with respect to light travelling through regions without the phase shifter. During the exposure, the electric fields diffracted underneath an absorber from two neighbouring apertures are then have out of phase by π. The fields interfere destructively, improving the intensity profile to allow improved resolutions to be obtained. The alternating phase-shift mask is a common technique used by the semiconductor industry to obtain a resolution enhancement.

1.3.4 Conformable Contact Photolithography

Conformable contact photolithography is a variant of the contact lithography described in Section 1.1.1 with the important difference being the conformability of the mask. The first work using conformable photolithography masks was reported by Smith in 1969 for the fabrication of surface acoustic wave devices [42]. In this initial work, 0.25 mm thin glass was used as the mask substrate and an amplitude mask pattern was formed in chrome via a 10× reduction using a photo-repeater. The thinness of the glass compared to the conventional 1.5 mm thick glass masters meant that the mask was able to flex to attain more uniform contact over large areas of the substrate during the contact print. Contact printing without conformable contact is a misnomer as true contact is not attained, technically it is proximity printing. In Smith's initial work, mechanical pressure was used to maintain contact and surface acoustic wave devices with interdigital electrode periods of 2.5 μm were replicated.

In later work Smith demonstrated 400 nm linewidths on an 800 nm pitch using the same technique but this time using vacuum pressure to attain intimate contact between the mask and substrate [43]. A distinct advantage of this system, he notes, is the insensitivity of the linewidth to exposure time when operating in true intimate contact, which is important in the manufacture of surface acoustic wave devices whose performance varies with changes in the linewidth. In this paper Smith hints at the prospect of achieving higher resolutions — "Although no effort has been made to explore the ultimate limitations of the conformable mask photolithography, our observations lead

us to speculate that narrower linewidths than reported here are probably possible, but may require the exercise of finer control over photoresist thickness and exposure parameters". To put this in the context of the time, the IC industry believed that 1 μm linewidths were the limit of contact printing. Smith's prophetic words turned out to be quite accurate, and in the following sections, sub-100 nm resolutions are reported with conformable masks.

A further important advantage of the conformable mask that is cited is the absence of damage or wear through repeated use [44]. Conformable masks are able to conform around topography created by contaminants such as dust particles, preventing large stresses developing at these points. In comparison, conventional photomasks used in contact lithography are rigid, so that when contaminants are trapped between the mask and rigid substrate, high stresses develop which can degrade the mask.

Recently, there has been renewed interest in conformable contact lithography with a number of different approaches used. Conformable substrates have been proposed as an alternative to the conformable masks previously described. Some innovative variations of the conformable mask have also been developed; the embedded amplitude mask, the light-coupling masks and the membrane masks used in this thesis work for evanescent near-field optical lithography. The three types of conformal mask are depicted in Fig. 1.7 for comparison. Various conformal contact schemes are described in the following sections, along with reported results.

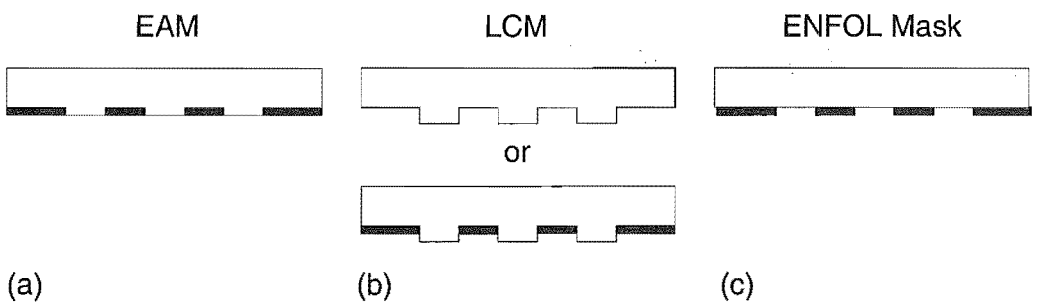


Figure 1.7 Comparison of different mask types used for conformable contact lithography (a) an embedded amplitude mask, (b) a light-coupling mask, and (c) an ENFOL mask.

1.3.4.1 Conformable Substrates

An alternative to using a conformable mask with a rigid substrate is to use a rigid mask with a conformable substrate. Ono *et al.* have replicated gratings with 500 nm periods and large duty cycles, *i.e.*, masks with large linewidths and small apertures (~150 nm) [45]. They formed substrates from 1 μm silicon diaphragms and performed a contact exposure with a standard chrome on glass mask using vacuum pressure and a mercury arc lamp source. They investigated the effect of polarisation on their results,

and claim that by reducing the far-field transmission by using a linearly polarised source (transverse electric), improved results are obtained. Further results using this technique have not been reported.

1.3.4.2 Conformable Embedded-Amplitude Masks

A modification of the conventional amplitude mask proposed by Goodberlet, is the conformable embedded-amplitude mask (EAM) [46]. The absorber is embedded into a thin transparent mask substrate resulting in a flat mask profile as shown in Fig. 1.7(a). The EAM consists of a fused silica substrate, 150 μm thick, with Cr absorbers, embedded into etched silica regions via a trilayer resist process. There are a number of advantages of an embedded absorber. A flat mask is expected to reduce local deformations around the absorber areas, to reduce pattern-displacement errors. An average pattern-placement error of 58 nm was measured for an exposure with the EAM over an area of $2\,\mathrm{cm}^2$ [46]. An additional advantage of the EAM is that the regions between the absorber have a higher refractive index than air or vacuum, these act to enhance the patterning resolution by taking advantage of the substrate's refractive index. Goodberlet *et al.* also suggests that the planarity of the mask may protect the absorber and minimise particle contamination. The EAM mask has been used to reproduce 100 nm line and space structures, using a source with a wavelength of 220 nm [46,47]. Attempts at multi-level alignment are also in progress [48].

1.3.4.3 Light-Coupling Masks

Light-coupling masks (LCM) for conformable contact photolithography have been developed by a group at IBM's Research Laboratory in Zurich [49, 50]. The masks are made of an elastomeric material, similar to that developed by Whitesides for microcontact printing and allow for some degree of conformal contact. The mask is not a conventional amplitude mask with absorbing and transparent regions, instead it consists of a surface relief pattern.

The light-coupling mask (LCM) has different contrast mechanisms depending on the size of the features d, relative to the exposing wavelength λ [49, 50]. Two forms of the LCM are discussed, one with and one without an absorber layer; both LCMs are illustrated in Fig. 1.7(b). For large features ($\lambda \ll d$), the LCM without the absorber creates some contrast owing to a reduction in intensity underneath the gap because of reflective losses at the mask/air and air/substrate boundaries. This contrast mechanism is generally not high enough for lithography and an absorber deposited onto the recessed regions improves the contrast. This solution places the absorber an additional distance from the photoresist, this is counter-intuitive in terms of near-field imaging, but for large features being in the near field is not so imperative. The LCM without the absorber appears like a binary phase mask as described in Section 1.3.3. This introduces an

additional contrast mechanism when the depth of the LCM protrusion gives a phase change near an odd multiple of π. For the regime where $\lambda \leq d$ and using no absorber, optical modes are excited in the protruding region of the LCM mask that can be visualised in the photoresist [50]. The final regime, where $\lambda > d$, a wave-guiding effect focuses the light into the higher refractive index material, amplifying the intensity below the protruding mask regions to provide contrast in the photoresist without the need for an absorber.

The smallest results reported with the LCM are 190 nm-period gratings formed using an LCM with a low surface relief pattern (70 nm deep). An exposure wavelength of 256 nm was used and no absorber was applied [50]. In this case the low surface relief means that the phase difference isn't near π so the destructive interference isn't significant. No further work has been reported in this area.

1.3.4.4 Evanescent Near Field Optical Lithography

Evanescent near-field optical Lithography (ENFOL), the subject of this thesis, is another conformable contact lithography technique. The major difference between ENFOL and the traditional contact lithography is the diffracted optical region that the image of the mask is collected from. While both operate in the near field, ENFOL is able to operate where the contrast of the intensity patterns is a result of electromagnetic waves that are evanescent, *i.e.,* they decay exponentially in the transmission direction. Traditionally contact lithography has relied on propagating electromagnetic waves to obtain exposure contrast. The short-range nature of the evanescent fields (in the order of tens of nanometres when operating at UV wavelengths), make it essential that the photoresist coated substrate is sufficiently close to the mask to take advantage of the image contrast in this region. Membrane masks ($\sim 2\,\mu$m thick) have been developed that are flexible and able to conform over topography. Thin photoresist layers are also required so that the image capture occurs in the evanescent near field. The style of the mask shown in Fig. 1.7(c) is similar to the standard contact masks with the absorber patterned on top of the mask substrate. During an exposure, the mask is held in contact with the substrate using vacuum pressure. Figure 1.8 is a schematic of the ENFOL technique.

The crucial difference of where the image plane is located with respect to the mask during an exposure, has an enormous impact on the resolution that can be attained. By operating in the evanescent near field, resolution better than the conventionally considered diffraction limit with optical projection lithography can be achieved. In associated work, resolution of 70 nm lines on a 140 nm period have been demonstrated with a mercury arc lamp, with actinic exposure wavelengths of 365–436nm [51–53]. The exposures were made into a thin single-layer resist and etched into silicon. Simulations performed by the author also indicate the tantalising prospect of image contrast for

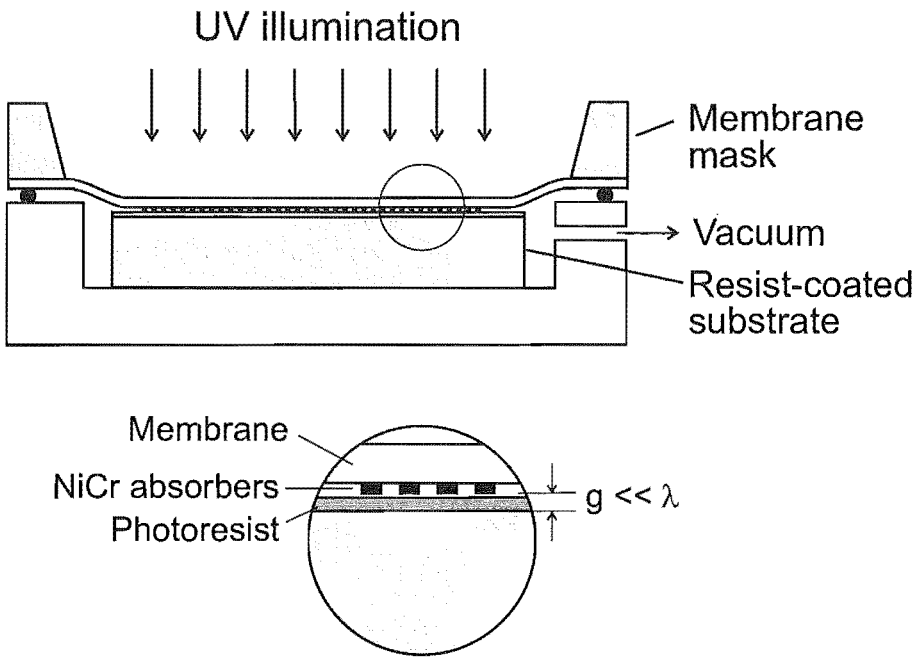


Figure 1.8 Schematic of ENFOL technique.

gratings with periods of $\lambda/20$ [54]. This is a key result discussed further in Chapter 6 of this thesis.

1.4 THESIS OUTLINE

The remainder of this thesis is structured as follows. Chapter 2 provides some background to optical resolution and defines a benchmark for the diffraction limit with respect to lens-based optical systems. The formulae considered to define the resolution limits for contact and proximity photolithography techniques are also stated. An introduction to the evanescent near field is presented as well as scanning near-field optical microscopy, a related technique achieving high resolution by utilising evanescent near field information. Basic grating theory is presented as background for some of the analysis presented in the simulation work in future chapters.

In Chapter 3 the main experimental techniques used in this thesis are outlined. Chapter 4 describes the fabrication of conformable membrane masks, a key component for an evanescent exposure. Flaws in the mask making process are discussed and possible solutions are proposed. In Chapter 5 experimental results are presented of some evanescent near-field optical exposures into thick resist. These highlight some of the key characteristics of the ENFOL exposure. The development of an additive pattern transfer process is described that uses a trilayer resist system for lift-off metallisation. A number of trilayer resist systems are trialed and an absorbing dye-doped bottom layer is developed to reduce substrate reflections.

Chapter 6 reports the results of simulation work that investigates what the fundamental limit of a technique such as evanescent near-field lithography might be. A rigorous electromagnetic technique, the multiple multipole program, is used to evaluate the intensities behind sub-wavelength grating structures. The decreasing depth of field as the grating period is reduced is analysed. Other grating parameters such as duty cycle, the radius of curvature of the conductor corners, conductor thickness and conductor material are varied with a view to optimising ENFOL, and to investigate process latitude.

In Chapter 7 further simulation results are presented that investigate the transition from propagating to evanescent exposures for grating structures. An interesting phenomenom is noted at the boundary of these two regimes where the effective illumination wavelength is equal to the grating period. In this vicinity a period halving can be observed in the TM polarisation due to an interference effect. This observation has lead to the proposal of a novel evanescent interference lithography technique that obtains an enhanced period halving in the near field when the exposure is made at a wavelength coinciding with a grating resonance.

In Chapter 8 the fabrication of a sub-wavelength wire-grid polariser is presented. It acts as a proof of concept device, being fabricated using an ENFOL exposure, and the pattern transferred using the developed lift-off metallisation with the trilayer resist system previously described in Chapter 5. Measurements of the polarisers' extinction coefficients are presented and compared with simulated results. Finally in Chapter 9, the conclusions of this thesis work are presented and directions for future work outlined.

Chapter 2

RESOLUTION AND THE EVANESCENT NEAR FIELD

2.1 INTRODUCTION

This background chapter discusses some of the important concepts addressed in this thesis, namely resolution in an optical imaging system and the evanescent near field. The diffraction limit is defined, so that it may be used as a benchmark to compare the resolution attainable with ENFOL. Previously considered resolution limits of contact lithography are also discussed.

Scanning near-field optical microscopy (SNOM) is briefly reviewed as it provides a pertinent example of the resolution attainable in the near-field via the use of evanescent fields. SNOM is one of a number of new microscopy techniques, including scanning tunnelling microscopy (STM) and atomic force microscopy (AFM), that have recently been developed. While the contrast mechanisms are different for each technique, they share in common the importance of having the detection probe in close proximity with the sample. As the detection mechanism for SNOM is similar to the exposure mechanism for ENFOL, some interesting parallels in resolution can be drawn.

Concepts of the near and far field are reviewed and the meaning of the evanescent near field is clarified. As resolution is defined in the context of grating structures for the experimental and simulation work presented here, some grating theory is briefly presented as background for the analyses in Chapter 6 and Chapter 7.

2.2 RESOLUTION LIMITS

The resolutions attainable by optical systems are fundamentally limited by diffraction, that is the spreading or bending of the electromagnetic wave as it passes through apertures, or around obstacles. Resolution in an imaging system can be defined as the ability to distinguish between two closely separated sources or objects. Clearly this is a rather subjective definition as "the ability to distinguish" does not specify a quantifiable threshold. Two widely accepted, yet disparate resolution criteria for

astronomical imaging and microscopy are the Rayleigh and Sparrow criteria [55]; these quantify when two point sources with Airy disk intensity distributions can be resolved. The Rayleigh criterion states that two incoherent point sources are just resolvable when the central maximum intensity of one overlaps the first intensity null of the second. This equates to a minimum spacing l_{min} between the two objects of

$$l_{min} = 0.61\frac{\lambda}{\mathrm{NA}} \tag{2.1}$$

where λ is the optical wavelength, and NA defines the ability of the optical lenses to collect the diffracted light and is given by $\mathrm{NA} = n\sin\theta$, where n is the refractive index of the surrounding material, and θ half the angle subtended by the lens aperture. The Rayleigh limit is illustrated in Fig. 2.1(a). Note that a small dip is present in between the two intensity maxima.

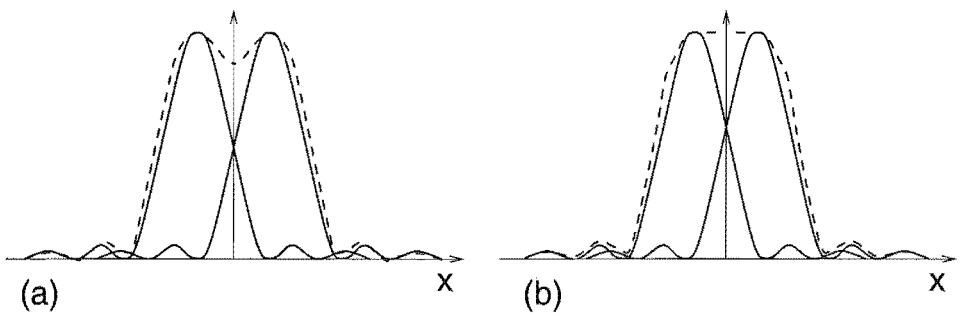


Figure 2.1 Resolution criteria for two point sources, (a) Rayleigh and (b) Sparrow. The solid lines represent the intensity function for each point source, and the dashed line, the resultant intensity.

While the Rayleigh criterion provides a convenient measure of resolution, a perhaps stricter, more appropriate definition is provided by the Sparrow criterion. The Sparrow criteria states that the resolution of two sources becomes unresolvable once

$$\frac{\partial^2}{\partial x^2}[I(x)]\,\Big|_{x=0} = 0, \tag{2.2}$$

where $I(x)$ represents the intensity as a function of x. This is the equivalent of starting with the Rayleigh criteria and shifting the sources closer together until no dip in the intensity exists. This can be considered the diffraction limit for two point sources, and equates to a separation given by

$$l_{min} = 0.47\frac{\lambda}{\mathrm{NA}}. \tag{2.3}$$

This resolution criterion is illustrated in Fig. 2.1(b).

The Rayleigh and Sparrow criteria have limited relevance for lithography. Firstly, note that as they stand they represent an optimistic diffraction limit, the intensity

curves illustrated in Fig. 2.1 would be intolerable for lithography. The criteria are also posed for self-luminous point-sources, whereas in lithography, the objects have an arbitrary shape and are back-illuminated. Another important point to keep in mind is the detection mechanism of the system; for astronomical imaging and microscopy it is the visibility as detected by the eye, whereas for lithography it is the photoresist's response to the exposure and subsequent development that defines the threshold for distinguishability. Nonetheless for optical projection lithography, the equation defining resolution has the same form as the Rayleigh and Sparrow criteria, but is modified to

$$l_{min} = \frac{k_1 \lambda}{\mathrm{NA}}. \tag{2.4}$$

The multiplier in the resolution formula of Equ. 2.1 and Equ. 2.3 is replaced by k_1, a factor which is an amalgamated process parameter that combines the effect of resolution enhancements in the exposure as well as incorporating the resist response to the resultant field. Resolution enhancements include the use of different illumination schemes, such as off-axis illumination, to reduce the k_1 factor. More sophisticated mask technology, such as alternating phase-shifting masks briefly described in Section 1.3.3, can also improve the pattern contrast and lower k_1. Typically k_1 is between 0.5 and 1 for standard manufacturing [13].

All of the previous resolution formulae specify the resolution in terms of the spacing between two isolated objects. If instead a grating structure is considered, then the minimum resolvable grating period p_{min} can be expressed as

$$p_{min} = \frac{\lambda}{\mathrm{NA}} \tag{2.5}$$

for normally incident illumination, and

$$p_{min} = \frac{\lambda}{2\,\mathrm{NA}} \tag{2.6}$$

for obliquely incident illumination. The p_{min} defined for oblique incidence defines the best possible resolution obtainable with propagating waves. It is consistent with an exposure with obliquely incident light, where a minimum of the zeroth and a first order beam is collected. The advantage of illuminating with obliquely incident light is illustrated in Fig. 2.2. With obliquely incident light, the first diffracted order can be diffracted at a larger angle α (corresponding to diffraction from a grating with a smaller period), and still be collected by the lens, compared with normally incident light. By shifting the zeroth propagating order off-axis and collecting only one of the ± 1 diffracted orders increases the effective NA of the lens. Note that Equ. 2.6 also corresponds to the theoretical limit of k_1 equal to 0.25 once $p_{min} = 2l_{min}$ is substituted. Assuming the best case scenario of a NA equal to one[1] the best resolution possible is

[1] An NA of 1 assumes that the optical system is not submerged in a high refractive index material.

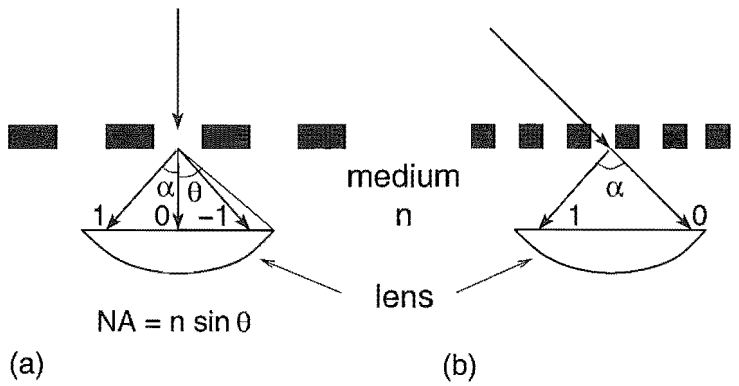


Figure 2.2 Schematic illustrating the advantages of off-axis illumination, (a) normal illumination, (b) oblique incidence. The numerical aperture (NA) is also defined for a lens in a medium of refractive index n.

$p = \lambda/2$.

The resolution limits described above all relate to lens-based optical systems. For contact and proximity lithography schemes, the theoretical resolution is given by Bowden [56] as

$$p = 3\sqrt{\lambda(s + (1/2)z)} \tag{2.7}$$

where p is the grating period, s the gap between the mask and photoresist surface, and z the photoresist thickness. This is derived from the Fresnel equations which make the standard Fresnel assumptions that the distance from the diffracting object to the image plane is large compared to the dimensions of the diffracting aperture. For $s >> z$, this simplifies to the familiar form $p \approx \sqrt{\lambda s}$. Equation 2.7 is inappropriate for describing the very near field. Applying this formula for a contact lithography setup similar to ENFOL, $\lambda = 436$ nm, $z = 60$ nm and $s = 0$ suggests a minimum period of 343 nm. Resolution of far smaller periods than this have been demonstrated by ENFOL [51–53] which indicate more rigorous simulation methods are required when working in the evanescent near field. Scanning near-field optical microscopy is a further example of an optical system achieving resolution beyond the expected limits.

2.3 SCANNING NEAR-FIELD OPTICAL MICROSCOPY

The advent of scanning near-field optical microscopy (SNOM) in the mid-1980's revolutionised conceptions of resolution in the optical regime. Until this time, conventional lens-based microscopy was diffraction limited to resolutions in the order of $\lambda/2$. Today resolutions of better than $\lambda/40$ have been demonstrated using SNOM [57]. The improved resolution capability is due to a radical change in approach that relies on short-range evanescent waves to provide information on a sub-subwavelength scale. An

optical fibre probe with an aperture typically tens of nanometres in diameter is scanned across a sample. There are many different configurations for SNOM, however the main modes of operation are collection mode (C-mode) and illumination mode (I-mode). These are illustrated in Fig. 2.3.

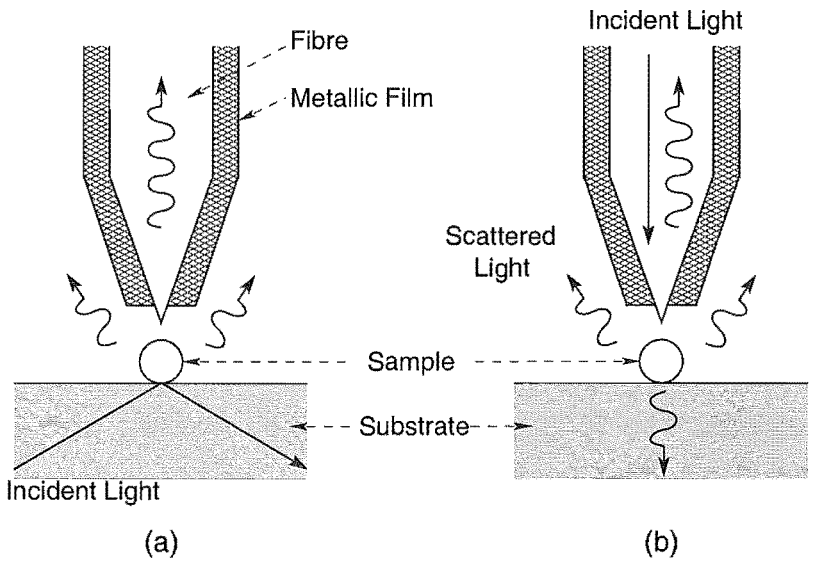


Figure 2.3 Schematic illustrating SNOM operating in (a) Collection mode (C-mode), (b) Illumination mode (I-mode) [58].

The C-mode SNOM relies on external illumination at an angle greater than the critical angle for total internal reflection, through an optically dense dielectric. When this condition is satisfied, evanescent waves are generated that are bound to the dielectric/air surface, which are then scattered by the sample. The optical fibre probe collects some of the scattered light, coupling the evanescent mode into a propagating mode that is guided by the fibre probe and then converted to an electronic signal via a photomultiplier.

In I-mode the sample is illuminated by light from the optical fibre. As the aperture of the probe is generally much smaller than the illumination wavelength, the sample is illuminated by evanescent fields. The sample scatters the field, some of which is collected by the probe in a similar way to the C-mode SNOM. Alternatively, other I-mode SNOM's use conventional collection optics to collect the scattered light.

While SNOM is still a diffraction-limited technique, the resultant resolution is predominately limited by the aperture of the probe and the sample-probe separation. As the evanescent fields are short-range, the sample probe separation is an important parameter for improving image contrast. Often an auxiliary method is applied to maintain a constant sample-probe distance such as shear-force feedback [59].

As the basic principles that govern resolution for optical microscopy also govern

those of optical lithography, it is realistic to expect similar improvements in resolution for optical lithography when operating in the near field. SNOMs configured for direct-write lithography have demonstrated 100 nm scale resolution with 450 nm illumination in I-mode and uncoated fibre tips. Smolyaninov *et al.* [60] used a scanned exposure into conventional photoresist while Madsen *et al.* [61] used the exposure from the fibre tip to locally oxidise a hydrogen-passivated amorphous silicon surface. The oxidised area then acts as an etch mask to a KOH wet-etch. While these techniques demonstrate the potential for high resolution for near-field optical lithography they still have the disadvantage of being slow due to their serial nature as do other scanned probe lithography techniques such as STM and AFM lithography [62], [63]. For years direct-write electron beam lithography has been demonstrating equivalent resolutions and better, with perhaps the only disadvantage compared to the scanned probe lithographies being the necessity to work under vacuum. To improve the feasibility of near-field optical lithography, ideally sub-diffraction limited resolution would be obtained with a parallel flood exposure. This is the aim of the evanescent near-field optical lithography (ENFOL) technique described in this thesis.

2.4 DIFFRACTIVE ZONES AND THE EVANESCENT NEAR FIELD

The resolution capability of lithography schemes such as ENFOL is limited by diffraction in the near field region, and more specifically by the range of evanescent fields. To optimise the performance of near field schemes, some understanding of the different diffractive zones is helpful. The region behind a diffracting object can be divided into two zones, a far field zone and a near field zone. The zones can be differentiated by the shape of the wavefront. Assuming an incident plane wave on a diffracting aperture, the far field is defined as the region where the waves have travelled enough distance from the diffracting aperture that they look almost plane again. In contrast, in the near field the wavefronts have significant curvature as they diverge from the aperture. The position of the boundary between the two zones is a function of the size of the aperture, with respect to the wavelength, through which the light is diffracted. This is illustrated in Figure 2.4.

Optics literature has long referred to the Fraunhofer far field, and the Fresnel near field diffracted zones. Incorporated in this terminology are also a set of approximations that encompasses the validity of approximations made in determining the fields in these regions. For the purpose of this thesis, the near field is separated into the evanescent near field and the Fresnel near field. The distinction is not a separation into two spatially adjacent regions, but more a distinction defining what assumptions have been made. The evanescent near field is immediately adjacent to the diffracting aperture. Here the field has a significant contribution from electromagnetic waves

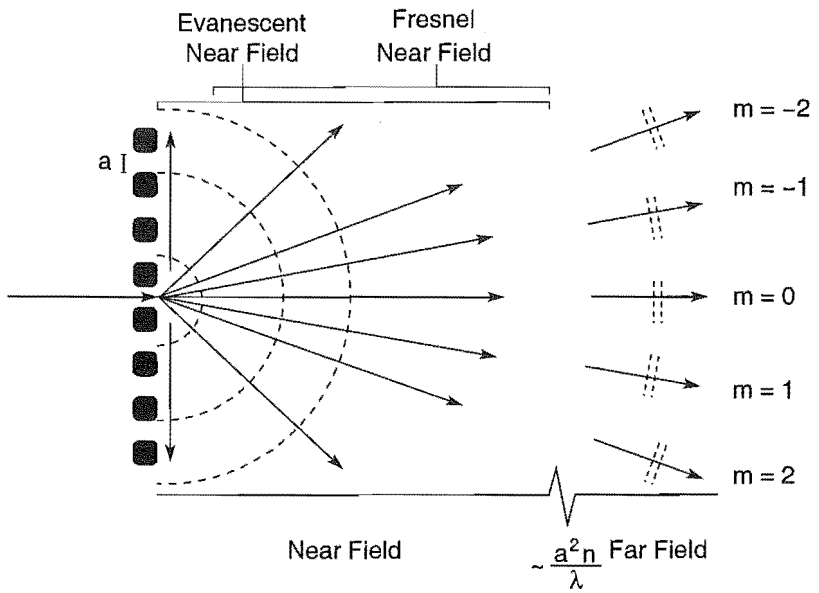


Figure 2.4 Diagram depicting the different near and far field zones, for transmission of a plane wave through a diffraction grating in a medium with refractive index n.

that are evanescent, *i.e.*, fields that have a characteristic exponential decay in the forward transmission direction. These evanescent waves are actually propagating surface waves, tightly bound to the diffracting object's surface. Their presence is accounted for in Maxwell's wave equations, but overlooked in less-rigorous, scalar treatments.

The Fresnel near field is also the region close to the diffracting aperture (but not immediately adjacent). It assumes scalar fields and hence its validity is improved for apertures that are large with respect to the wavelength. A further simplifying Fresnel approximation is that the distance from the diffracting aperture is large in comparison to the aperture dimensions [64], invalidating results in the immediate vicinity of the diffracting aperture.

Figure 2.4 illustrates the different diffraction zones for a transmission grating, illuminated with a plane wave. The light is diffracted into different orders m, that represent different spatial frequencies. To obtain the best representation of the object in a far-field reconstructed image, as many of these orders should be collected as possible. For lens-based optical equipment such as microscopes and optical projection lithography, only the light that reaches the far field is collected. The evanescent fields are not detected owing to their short-range. Light that is diffracted at high angles is also not collected further limiting the resolution obtainable.

2.5 SURFACE PLASMONS AND EVANESCENT WAVES

Surface plasmons (also known as surface plasmon polaritons) are closely associated with evanescent waves. Surface plasmons are oscillations or fluctuations of surface electron density. These charge fluctuations are coupled with electromagnetic surface waves that have their intensity maxima in the surface and exponentially decaying fields perpendicular to it, *i.e.,* evanescent waves. Surface plasmons are generated by a dynamic charge accumulation process or "sloshing" and are evident at the interface between a highly conductive metal (with a negative real part of the permittivity) and a dielectric. Surface plasmons have gained a high profile in the research literature recently following the paper by Ebbesen *et al.* in *Nature* [65], reporting unusually high zero order transmission spectra at wavelengths larger than the period of the hole array. The high transmission is credited to surface plasmon resonances that are then coupled into a propagating wave by scattering in the holes. Comparison of transmission through the hole arrays are made with those predicted by Bethe's [66], and show greatly enhanced transmission. The comparison should however be made, keeping in mind that simplifying assumptions are made by Bethe, such as uniform field hole aperture, that do not take into account the effect of surface plasmons and the coupled evanescent waves. While the "anomalous" or extraordinary transmission of light is discussed, Treacy reports with reference to gratings, that the transmission is correctly described by Maxwells's equations with the appropriate boundary conditions [67], and there is reason to believe this will hold for hole arrays also.

2.6 GRATING THEORY

Gratings are the structures analysed in the most detail in both the experimental and simulation work presented in this thesis. They provide a useful measure of resolution in terms of the smallest period that can be resolved with the least ambiguity. They also have well defined, distinct spatial frequencies (orders), providing a well understood framework for establishing the presence of evanescent waves. This section summarises some basic grating theory and derives an expression that defines when the orders diffracted from a grating change from propagating to evanescent [68].

To begin with, let us consider a transmission grating in a co-ordinate system as defined in Fig. 2.5. The grating equation

$$n \sin\theta_m = n \sin\theta_i + \frac{m\lambda}{p}, \qquad m = 0,\ \pm 1,\ \pm 2,\ \cdots, \tag{2.8}$$

defines at what angles θ_m the diffracted orders m are diffracted, for light incident with an angle θ_i, on a grating of period p, in a medium with refractive index n. For our purposes we will limit the analysis to illumination by light normally incident, $\theta_i = 0$ so

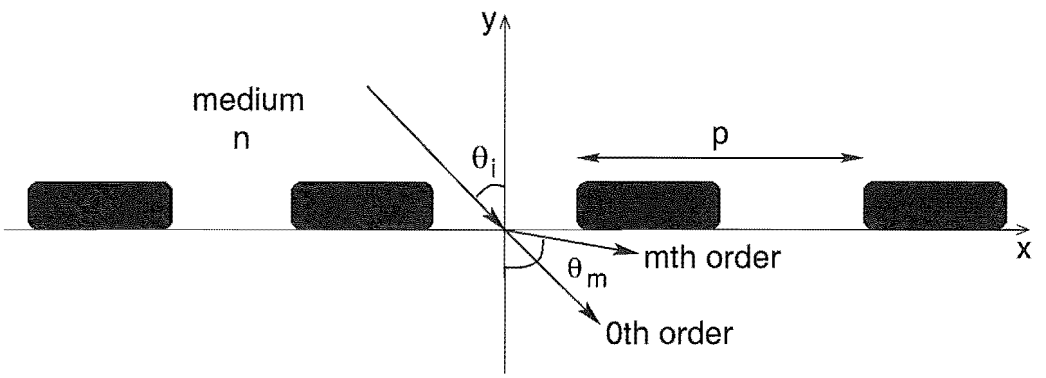


Figure 2.5 Schematic representation showing light diffraction by a transmission grating, in a classical diffraction mounting. The coordinate system is introduced and angles of incidence and diffraction.

Equ. 2.8 simplifies to

$$\sin\theta_m = \frac{m\lambda}{pn} \tag{2.9}$$

The grating equations of Equ. 2.8 and Equ. 2.9 can be satisfied by more than one value of m, and for all orders m that satisfy $|\sin\theta_m| \leq 1$ the diffracted wave of order m will be propagating, characterised by fields with a e^{iky} dependence. From the wave equation,

$$k_x^2 + k_y^2 = k^2 = \left(\frac{2\pi n}{\lambda}\right)^2 \tag{2.10}$$

where k_x and k_y are the components of the wavevector in the x and y directions respectively. Note, $k_z = 0$ as we assume that the plane of incidence is perpendicular to the grating; this case is called the classical diffraction mounting. For a grating illuminated at normal incidence, k_{mx} is given by

$$k_{mx} = \frac{2\pi m}{p}. \tag{2.11}$$

Solving for the vertical wavevector component k_{my} is then,

$$k_{my} = 2\pi\sqrt{\left(\frac{n}{\lambda}\right)^2 - \left(\frac{m}{p}\right)^2}. \tag{2.12}$$

It is clear from Equ. 2.9 and Equ. 2.12 that for the condition $|\sin\theta_m| > 1$, the term under the square root becomes negative and k_{my} becomes imaginary. This is equivalent to a field that decays with distance in the transmission direction y, characteristic of evanescent waves.

While the evanescent orders decay exponentially in the transmission direction, they

propagate in the x-direction as surface waves tightly bound to the grating surface. It is the presence of these evanescent waves that we credit for the high resolution obtained with our ENFOL technique. In summary, when the period p is

$$p < m\lambda/n \tag{2.13}$$

the mth orders will be evanescent. If we now choose $p < \lambda/n$, we have a totally evanescent exposure with only the $m = 0$ order propagating, and all other orders evanescent. This case also equates to the resolution limit for gratings for lens-based optical systems, with normally incident light as given in Equ. 2.5. An image reconstructed from far-field information would not reconstruct an intensity representing the grating, as the orders that provide intensity contrast would have decayed away. Only the $m = 0$ component would be collected, and this provides only a constant intensity, parallel to the grating (in the x-direction). In order to resolve the grating, the diffracted light should be collected before the evanescent waves have decayed away, *i.e.*, in the evanescent near field.

Grating theory provides useful information regarding the form of the diffracting orders, propagating or evanescent, as well as predicting the angles of the diffracting orders. No information regarding the relative intensities of the diffracted orders is provided. To determine this information, solutions can be found either analytically, with the application of simplifying assumptions, or by using rigorous electromagnetic simulation techniques employing numerical methods. There is also no substitute for experimental evidence.

Chapter 3

EXPERIMENTAL TECHNIQUES

To investigate the resolution limit of ENFOL experimentally, a number of auxiliary experimental techniques were used. This chapter discusses some of the main techniques of interest, such as electron beam lithography that was used to define the ENFOL masks, and reactive ion etching that was used for pattern transfer following the ENFOL exposure. To evaluate the results at these scales, high magnification microscopy techniques such as atomic force microscopy and scanning electron microscopy were required. Transmission optical spectroscopy was used for evaluating the transmission properties of various materials, such as membranes and resists.

As much of the equipment was new to the University of Canterbury, prescribed procedures, methods and recipes were often not in place. The author was involved in establishing experimental systems and methods, in particular for the electron beam lithography tool.

3.1 ELECTRON BEAM LITHOGRAPHY

Electron beam lithography (EBL) is the ideal tool for patterning the ENFOL masks. It is capable of high resolution which is essential if patterns on the mask are to be made small enough to adequately test the resolution of ENFOL. As EBL is a maskless technique, patterns can be easily compiled in software making the mask design process very flexible.

The EBL tool used in this work is a converted Philips PSEM500 scanning electron microscope. It was formerly owned by a group at Glasgow University who converted it to a lithography tool by the introduction of some electronics providing a pattern generator (with a vector scan), and a beam blanking system to turn the beam on and off. A schematic of the modified Philips system is illustrated in Fig. 3.1. It is comprised of the following components and systems:

The Electron Gun

The source of the electron gun is a heated tungsten hairpin from which thermionic emission is extracted by the application of a 1.5–50 kV voltage. The Wehnelt acts to

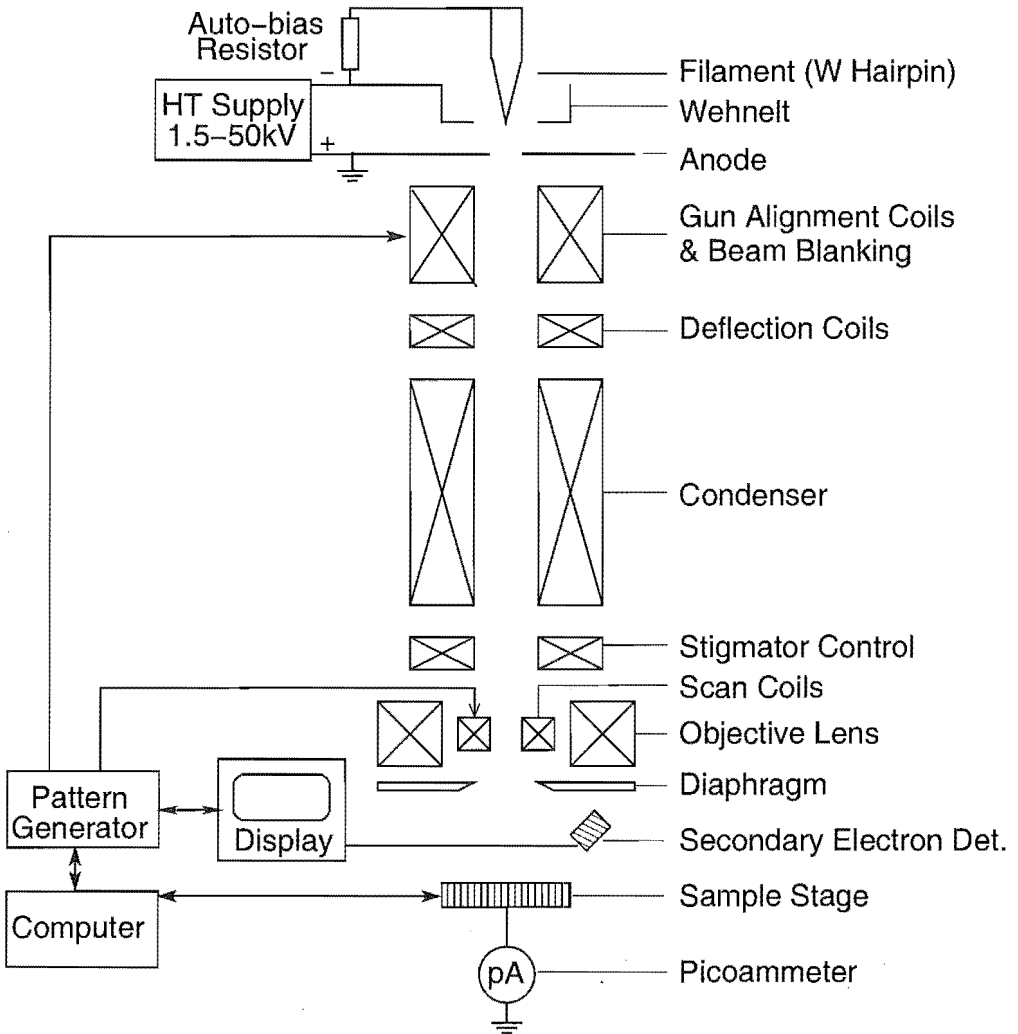


Figure 3.1 Schematic of the Philips PSEM 500 electron beam lithography tool.

shield the hairpin and is biased to a more negative voltage by use of an auto-biasing resistor. The auto-biasing resistor provides current feedback to stabilise the emitted current. The electrons are accelerated by the field between the cathode filament and the anode.

Electron Optics

A series of magnetic lenses (or coils) are used to align and focus the electron beam. The gun alignment coils control the beam tilt and the deflection coils control the beam translation in the X/Y direction. These must be aligned when changing to a new filament and minor adjustments may be required as the filament can change over time. Beam blanking was added at the gun alignment coils; when the beam blank is on, the beam is tilted out of the electron optics path and onto a diaphragm.

The condenser consists of magnetic lenses to demagnify the electron beam cross-

over point. It is responsible for adjusting the beam current. A stigmator adjusts the beam roundness. Scanning coils deflect the electron beam in order to raster scan a sample or in the case of our EBL tool to perform a vector scan. The angle of deflection is dependent on the magnification setting of the microscope. Focusing is implemented by the objective lens.

Detector System

The Philips SEM uses a secondary electron detector. It consists of a scintillator and photomultiplier combination that collects the secondary electrons, which are the electrons scattered with low energies ($\leq$ 50 eV). The detector has a positively biased collector grid which deflects the secondary electrons but not the back scattered electrons (electrons with energies $>$ 50 eV). The deflected electrons are then accelerated to a +10 kV biased scintillator where light quanta are generated then collected and amplified in the photomultiplier [69].

Vacuum System

High vacuum is required in the electron column to prevent premature scattering of the electrons before reaching the sample. The Philips has a two stage pumping process; a rotary pump or pre-vacuum pump initially brings the pressure down in the column to about 0.1 Torr, then the diffusion pump brings the pressure down to typically between 10^{-5} and 10^{-6} Torr. The rotary pump acts as the backing pump for the diffusion pump; it does not operate continuously but switches between pumping and a vacuum reservoir [70]. A separate rotary pump evacuates the interlock, the intermediate chamber between atmospheric pressure and the high vacuum column that is used to load samples.

Beam Current Measurement

To be able to write high resolution patterns with the modified SEM, strict control of the delivered electron dose is necessary. This requires accurate measurement of the beam current so that beam dwell times can be determined to control the electron dose. To do this a picoammeter has been connected to the sample stage and a Faraday cup machined into the sample holder. The Faraday cup is a cavity, shaped to maximise the collection of electrons when the beam is placed over the centre of its aperture. Secondary electron emission out of the cup is minimised to obtain as accurate a reading of the beam current as possible.

The Pattern Generator

The pattern generator controls the scan coils, beam blanking and the dwell time of the beam. The combination of the beam current and dwell time defines the primary exposure dose received by each pixel. Secondary exposure also occurs due to exposure from scattered electrons from pixels exposed close by, this secondary exposure is known as the proximity effect. The scan field is the area encompassed by the maximum deflection of the scan coils. It is divided into 4096 $\times$ 4096 individually addressable

pixels, which is fixed by the 12 bit digital to analog converter. The field size and hence pixel size varies with the magnification setting. Table 3.1 gives some guide to the scan field size and the associated pixel size obtained from experimental measurements. Altering the fine controls on the scan coil control unit can change the scan field size.

Table 3.1 Scan field size and pixel size versus magnification.

Magnification	X Field Size (μm)	Y Field Size (μm)	X Pixel Size (nm)	Y Pixel Size (nm)
40	3500	2250	860	550
80	1650	1100	400	270
160	870	600	210	150
320	430	300	110	72
640	220	150	54	37
1250	110	77	29	19
2500	58	39	14	9.5
5000	28	19	6.9	4.5
10000	14	10	3.5	2.4
20000	6.9	5	1.7	1.2

Electron Beam Scanning System

The electron beam scanning system (EBSS) is the software on the EBL system's computer that translates the generated pattern files into scanning patterns on the Philips Electron microscope. From EBSS the user can control the x-y translation stage, and set scans running. The beam current and magnification of the SEM must be manually entered, along with the pattern file to be read. While EBSS can operate in two modes, manual or automatic, only manual mode was utilised. This was because the pattern generator was prone to crashing and refocusing throughout an exposure run at regular intervals gave better results.

3.1.1 Defining the Pattern Files

Before performing an EBL exposure, pattern files must be created, the equivalent of a lithography mask. A primitive computer program for creating pattern files existed dating back to the mid-1980s but this was slow and cumbersome. The author wrote a C program 'cifconvert' to convert Caltech Intermediate Form ('.CIF' files) to a format compatible with the Philips' pattern generator. This meant pattern files could be created using an existing commercial integrated circuit layout tool (L-Edit) then designs exported in a '.CIF' file format followed by a conversion to the non-standard format required by the Electron Beam Scanning (EBSS) program. The process flow from designing the original pattern file to the actual scanned output is described in Fig. 3.2.

Using L-Edit's existing graphical interface simplifies the pattern drawing task, and

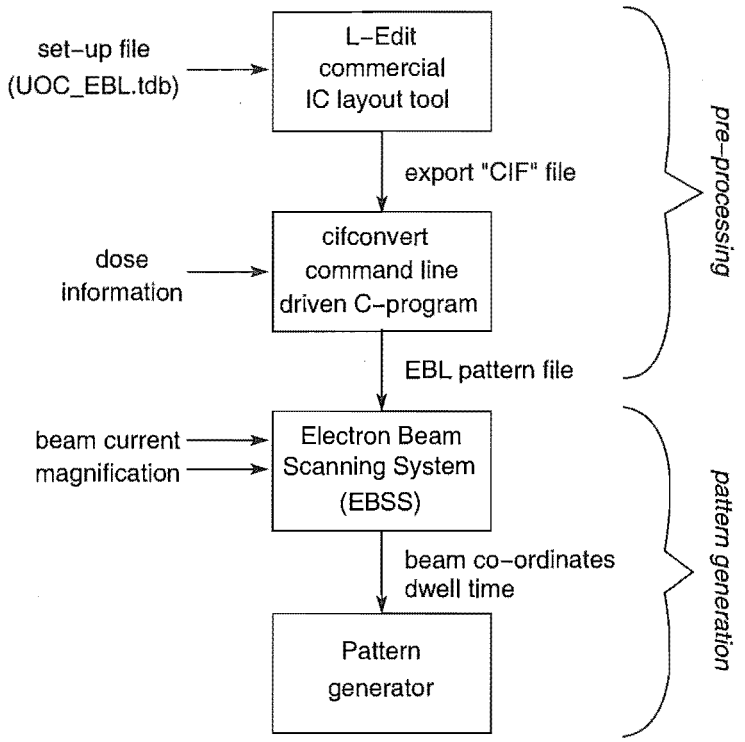


Figure 3.2 Flowchart indicating process from defining the pattern files through to generating the patterns.

saved programming a customised interface for producing EBL pattern files. To make the exported '.CIF' files compatible with the cifconvert program the user should execute the *Replace Setup* command and specify the 'UOC_EBL.tdb' file. This sets the layers and design rules correctly. The only valid drawing object is the box tool, all other object types are ignored. An internal unit (or lambda) is equivalent to one pixel. All patterns should be drawn within the 0 to 4095 x,y co-ordinate space in L-Edit. Different layers in L-Edit can be specified and these are interpreted as different doses by cifconvert. Once the desired pattern has been created in L-Edit this should be exported in a '.CIF' file format.

The cifconvert program is a command-line based program which should be executed from the DOS prompt. The command-line arguments are as follows –

`cifconvert` *`inputfile outputfile dose0 dose1 dose2 dose3 ...`*

where the *inputfile* is the '.CIF' format file, the *outputfile* is the resultant file to be used with EBSS and *dose0* is the exposure dose in μC/cm^2 for all boxes defined on layer 0, *dose1* is the exposure dose for all boxes defined on layer 1 for example. Writing a batch file to convert a number of files at once is useful as it also acts as a record of the files and the assigned doses. The cifconvert program implements a simple sort algorithm to minimise the amount of jumping around required of the beam to reduce beam skewing. Beam skewing results from hysteresis in the scan coils and from the finite time that

the beam takes to settle following beam unblanking. The user can have some control of the sorting by 'grouping' boxes together to force the sort of individual groups. The grouping is done in L-Edit using the 'Group' command.

Once suitable pattern files have been created, they are loaded onto the EBL systems' computer ready for patterning. The final step is the exposure of samples coated in an electron sensitive resist, typically poly(methyl methacrylate) (PMMA).

3.2 REACTIVE ION ETCHING

Reactive ion etching (RIE) is used for etching the trilayer resist stack that has been used as part of the additive pattern transfer process for ENFOL as described in Section 5.4. RIE demonstrates good control over etch profiles which is important in this application as anisotropic steep sidewalls, which are slightly undercut, are required for successful lift-off. The principle of RIE, or dry etching as it is also known, is the removal of material via a combination of physical bombardment and chemical reactions. The main advantage of RIE over wet etching techniques is the increased control over the etch. A number of parameters such as pressure, temperature and plasma power can be adjusted to control the anisotropy of an etch as well as various passivation mechanisms available through the choice of gas etchants. Being a dry technique there are fewer waste products to dispose of.

3.2.1 Principles of Reactive Ion Etching

Reactive ion etching is a plasma etch process that combines controlled energetic ion bombardment with chemically reactive interactions to achieve high selectivity and a highly anisotropic etch. There are a number of different configurations. The parallel plate electrode RIE is most common, and more recently high ion density reactors have become available such as electron cyclotron resonance (ECR) reactors, helicon and helical reactors, and inductively coupled plasma (ICP) reactors. They are all based on the same basic principles, however the high ion density reactors are configured to achieve higher ionisation densities compared to the parallel plate RIE. Higher densities of ions at lower pressures increases etch rates, while still maintaining a highly anisotropic etch.

A plasma is ignited by the application of radio frequency (RF) energy, typically at 13.56 MHz. This promotes ionisation of the constituent atoms and molecules in the process gases to create a plasma. This is essentially an electric discharge process. The plasma consists of reactive radicals, as well as a large proportion of neutral species. Of the reactive species, the majority of positively charged species are singly ionised atoms or molecules, while the majority of the negatively charged species are electrons. The plasma appears to glow owing to photon emission during electron recombination

processes, with colours characteristic of the different gases. The higher mobility electrons are able to diffuse more quickly through the plasma, recombining on the chamber walls. In the region immediately adjacent to the chamber surfaces, a sheath region forms. This is an area of low charge density created by the recombination. As the electrons are able to diffuse more quickly in the plasma, a higher concentration of positive charge remains in the plasma.

The sheath plays a major role in etching. Because it has a low charge density, large fields can develop that accelerate positive ions across the sheath and onto the sample. Figure 3.3 shows how the electrical potential is distributed in the chamber for a simple parallel plate electrode system. The cathode becomes negatively biased

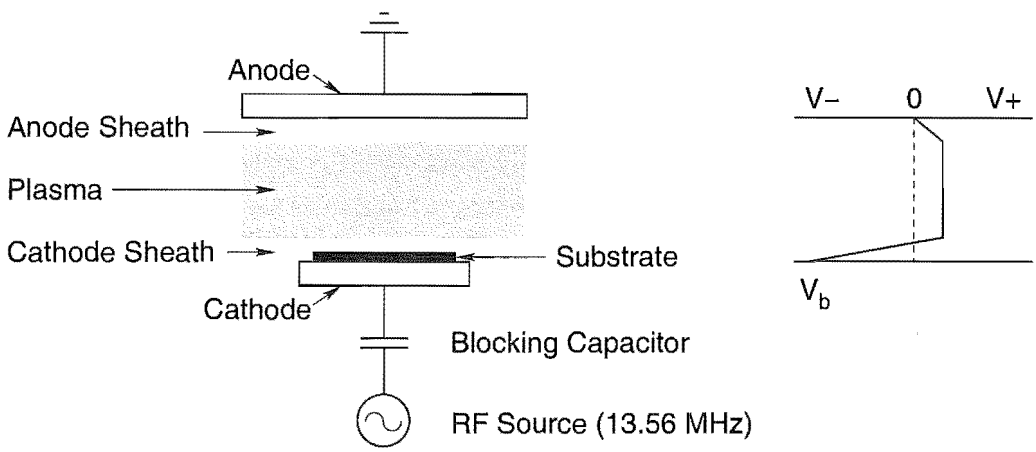


Figure 3.3 The electrical potential distribution across the discharge region.

during the etching process. This is a direct result of the higher mobility of the electrons over the positively charged ions. During the positive half of the RF cycle, electrons are attracted to the cathode, while on the negative half of the RF cycle, positive ions are attracted. Initially, more electrons impinge on the cathode than positive ions, causing the cathode to take on a negative bias, this in turn means that there is a reduced percentage of the cycle when the cathode is more positive than the plasma potential. A steady-state is reached, when the number of electrons equals the number of ions hitting the cathode. The DC voltage of the cathode at this steady state condition is called the self-bias voltage, or bias voltage (V_b), it gives important information as to the energy of the impinging ions. A V_b large enough to obtain an anisotropic etch is generally desirable [71].

A much larger voltage is apparent over the cathode sheath, compared to the anode sheath as shown in Fig. 3.3. This is a consequence of the anode area being much larger than the cathode, creating a larger capacitance and hence a lower voltage. This is desirable as larger fields, and hence more energetic bombardment occurs on the samples on the cathode.

3.2.2 Reactive Ion Etching Mechanisms

Reactive ion etching can be separated into four basic etching mechanisms [72], 1) sputtering, 2) chemical gasification, 3) energetic ion-enhanced chemistry and 4) inhibitor ion-enhanced chemistry. The dominance or otherwise of each of the different processes is dependent on a number of parameters, including the material to be etched, the process gases, pressure, RF power, frequency, temperature, and etcher configuration.

Sputtering is a low pressure process that creates a long mean free path for radicals. The radicals that cross the sheath are very energetic, and collide with the sample at predominately normal incidence due to the small number of collisions. The energetic nature of the collision causes surface material to be ejected. This is the least selective of the etching processes, as the etching is due to physical bombardment as opposed to a chemical process. Vertical etching can be achieved using this process.

Chemical gasification on the other hand is a chemical process. The plasma provides the reactive etchant species which chemically react with the substrate to be etched. It is a very selective process, sensitive to differences in bonds and the chemical consistency of the material. It is similar to wet-etching in that it results in isotropic profiles.

Energetic ion-enhanced chemistry is a combination of the above two mechanisms, energetic ions impinging on the surface damage it, increasing the reactivity to speed up the chemical reaction.

Inhibitor ion-enhanced chemistry works by having two species in the plasma, etchants and inhibitors. Normally the etchant and substrate would react and etch isotropically. However, the inhibitor forms a thin film on vertical surfaces that see little or no bombardment. This film acts as a barrier to etching, improving the anisotropy of the process.

The dry-etching for the multi-layer resist systems was a combination of the first three processes. Chemically selective reactive processes were combined with low pressures, to obtain a high V_b to improve the anisotropy.

3.2.3 Machine Description

RIE was performed using an Oxford Plasmalab 80 system. It has a basic parallel plate electrode configuration with a cryogenic sample stage with a temperature range of -150°C to 200°C. A schematic of this system is shown in Fig. 3.4. The RIE can be separated into four subsystems: the chamber where the etching takes place with its RF source, the gas-handling system, the vacuum system, and the controller.

In the aluminium chamber there are two parallel plate electrodes; the upper electrode or anode that is grounded, and the lower electrode or cathode that is the drive electrode supplied with RF energy (at 13.56 MHz). The cathode is 200 mm in diameter, while the upper anode encompasses most of the chamber lid, approximately 280 mm

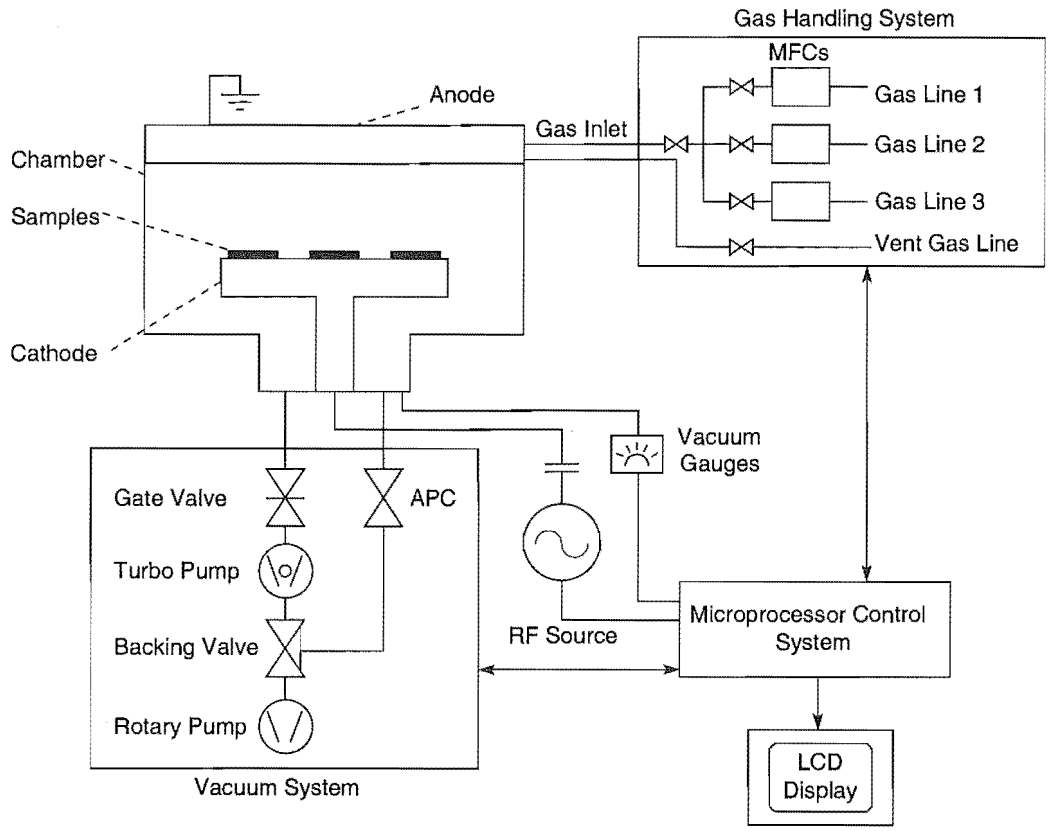


Figure 3.4 Schematic of the Oxford Plasmalab 80 Reactive Ion Etcher.

in diameter. The lower electrode's temperature can be controlled with liquid N_2 to cool the electrode down to -150°C and has a heating element to heat the electrode to a maximum of 200°C (these are not indicated in Fig. 3.4).

Process gases and a vent gas are supplied to the upper electrode. There are three process gas lines, allowing a maximum of three gases to be used simultaneously. In turn these three gas lines are connected in parallel to up to three gas cylinders. Available gases include SF_6, CHF_3, CH_4, O_2, N_2, and Ar. Mass flow controllers (MFCs) control the flow-rate through each of the gas lines. The vent line vents the chamber with nitrogen gas to bring the chamber back up to atmospheric pressure to allow loading and unloading of samples.

The vacuum system consists of a two stage pumping system to evacuate the chamber as well as to control process pressures during etching. Prior to etching, the chamber pressure is lowered to a base pressure, typically lower than 7×10^{-5} Torr, to reduce contaminant species in the chamber. A rotary pump acts as the roughing pump to lower the pressure from atmospheric, providing the first stage of vacuum. A turbomolecular pump then lowers the chamber pressure further and can achieve pressures of around 10^{-8} Torr. The rotary pump backs the turbomolecular pump's backing line. During the

etching process the gate valve is shut and all process gases are removed via the rotary pump, with the automatic pressure controller (APC) valve maintaining the process pressure. The chamber pressure is monitored via two gauges, a capacitance manometer gauge for measuring relatively high pressures such as during etching, and a Penning gauge for measuring low pressures such as the base pressure.

The control system is based on a microprocessor based control module. The user can specify pre-loaded recipes or set process parameters manually. The control module co-ordinates the operation of the vacuum system valves to control pumping and to set process pressures, operates the gas valves and MFCs for controlling the gas inputs, as well as controlling the temperature of the cathode, and the application of RF power.

3.3 TRANSMISSION OPTICAL SPECTROSCOPY

Transmission optical spectroscopy was used to measure the transmission of a number of materials over the course of this work. The transmission properties of the Si_xN_y membranes used for the ENFOL mask substrate were measured to help ascertain which wavelengths contribute to the ENFOL exposure. Experiments were also conducted to compare various dye-doped PMMA compositions and their effectiveness as an anti-reflection layer. The author evaluated a number of different options for an anti-reflection layer in the trilayer resist stack for pattern transfer experiments following ENFOL exposures.

The Cary Spectrophotometer 14 was used to perform transmission measurements on different films for assessing film absorption and to calculate refractive indices. A schematic of the spectrophotometer is illustrated in Fig. 3.5. It has two sources, a quartz halogen lamp that emits over a range of wavelengths from 3000 – 18000 Å and a deuterium lamp that operates from 2000 – 3000 Å. A double monochromator, consisting of a diffraction grating and prism, disperses the light to obtain a monochromatic output.

Transmission measurements can be taken over the range of wavelengths permitted by the source. The instrument has two light-tight chambers, so the sample to be measured can be compared with the reference sample to take into account any fluctuations in intensity, either temporally or with changes in wavelength. It has two modes of operation, manual and auto. In manual mode the slit is left at a fixed width; variation in the intensity will occur with wavelength due to the sources' non-uniform output as well as wavelength dependent losses in the optics. Auto mode adjusts the slit to try and maintain a fixed intensity for the reference.

Because the reference and sample beams were not calibrated satisfactorily to allow a single scan to compare the sample with a control sample in the reference beam path, data was collected using two scans. For the first scan, a control sample, for example a uncoated glass sample would be placed in front of the sample beam with nothing in the

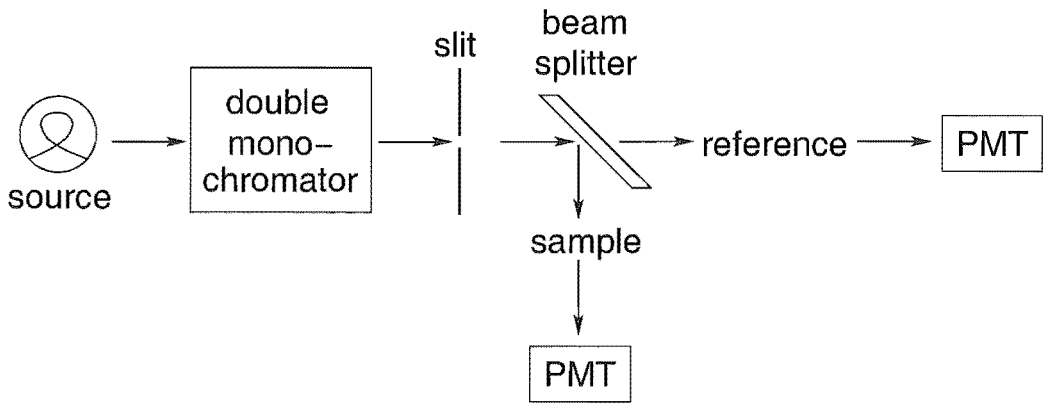


Figure 3.5 Schematic showing the principle of operation of the Cary Spectrophotometer 14.

reference beam. A scan would be conducted with a slit opening that would not saturate the PMTs for either beam. The second scan would be conducted with the sample in front of the sample beam and the reference unchanged. A fixed slit was used for both scans, a variable slit should also have been acceptable as each scan is referenced to the reference intensity, however by maintaining a fixed slit the reference intensity could be directly compared as a manual check. An advantage of using the auto mode with a variable slit would have meant that down at the lower wavelengths the slit would have been wider, providing a better signal to noise ratio.

3.4 ATOMIC FORCE MICROSCOPY

Atomic force microscopy (AFM) became available in the latter stages of this research with the purchase of a Digital Instruments Dimension 3100 (DI 3100) instrument. It was particularly useful for investigating the topography of resists, with little sample preparation required prior to imaging. Before the AFM became available, an SEM was used for the majority of imaging, but for this, insulating material such as resists, need to be treated with a conducting film to prevent charging. Small resist features were also susceptible to warping due to heating by the electron beam.

3.4.1 Principles of AFM

AFM belongs to the class of scanning probe microscopy techniques. An image is obtained by scanning a probe with a very sharp tip over a sample, with a tip to sample distance in the order of angstroms. The tip is attached to a cantilever which is deflected towards and away from the sample by separation-dependent atomic forces. In the common AFM configuration depicted in Fig. 3.6, the cantilever deflection is measured by a position-sensitive photodetector consisting of a laser diode and bi-cell split photodetec-

tor. From this deflection, an image of the sample topography can be obtained. There are many AFM configurations but for simplicity the configuration shown in Fig. 3.6 will be discussed, as this resembles that of the DI 3100 instrument [73].

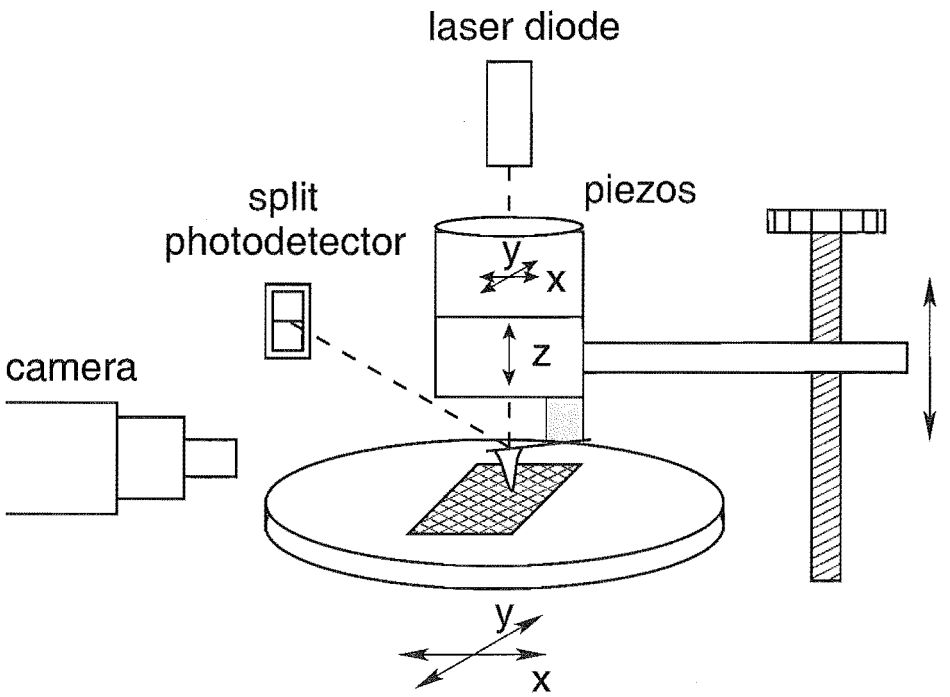


Figure 3.6 Schematic of an Atomic Force Microscope.

Coarse positioning of the sample in the x-y plane is provided by stepper motors, which shift the sample stage. An additional stepper motor controls the height of the probe, relative to the stage for coarse adjustments. A camera provides feedback to aid positioning of the sample and probe, so that specific regions of the sample can be located and scanned. When an image is to be captured, the tip is lowered onto the sample (at a distance dependent on the AFM mode), and the z piezoelectric scanner brings the tip to the final scan position. The x-y piezoelectric scanners move the cantilever in a raster scan, building up an image pixel by pixel. The maximum scan field is restricted by the range of the piezoelectric scanners, approximately 100 μm $\times$ 100 μm for the DI 3100.

The AFM instrument can be operated in a number of different modes: contact AFM, non-contact AFM, and tapping or intermittent-contact AFM. These modes correspond to operation at different probe/sample separations. Figure 3.7 illustrates the relationship between the distance of operation, the relative strengths of these forces and their direction [74].

In contact mode (repulsive mode) the tip makes soft physical contact with the

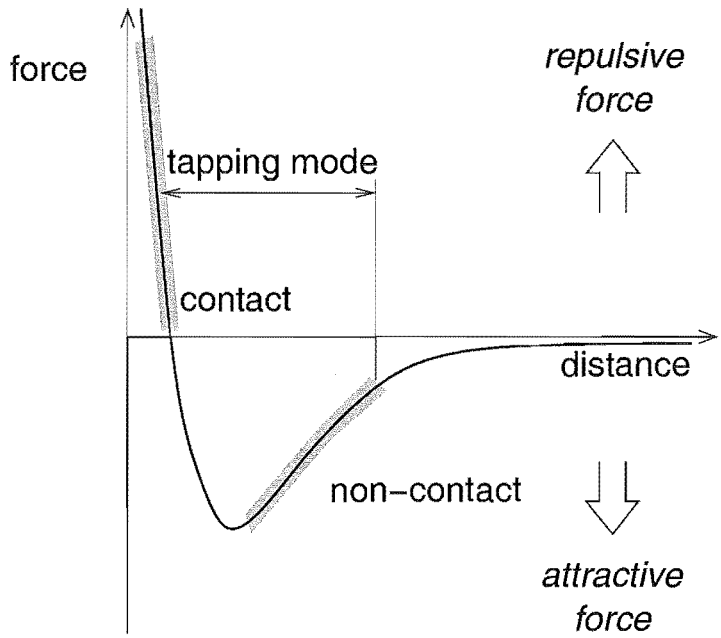


Figure 3.7 Graph showing the relationship between the AFM probe/sample separation and force for different AFM modes.

sample and a cantilever with a low spring constant is required. The advantage of operating in this mode is that the forces are large, and the distance-force curve quite steep so small deflections can be measured more easily. The spring constant of the cantilever should be lower than the effective spring constant of the surface to prevent surface deformation. During the scan the cantilever deflection is measured. The scan can proceed in either a constant-height or constant-force mode. In constant-height mode the deflection is translated directly into a surface topography, this is suitable for low relief surfaces. More common is the constant-force mode, which feeds back the deflection information to raise or lower the cantilever to maintain a constant force. The cantilever motion then reflects the surface topography.

In non-contact mode, no physical contact is made with the sample. The cantilever is oscillated at a frequency which is slightly above the cantilever's resonant frequency, just above the sample surface ($\approx$ 10–100 Å). The resonant frequency or vibrational amplitude is monitored, and kept constant by a feedback mechanism that shifts the cantilever vertically. It operates with small forces that are suitable for studying soft or elastic samples. Because these forces are small, they are also more difficult to measure to the same accuracy. A stiff cantilever is used to prevent the cantilever being pulled into contact with the sample. Non-contact mode requires a slower scan mode than contact or tapping mode AFM.

In tapping mode, an oscillating cantilever is brought close enough to the sample

that it just lightly taps the sample. It is less likely to damage soft samples compared to contact AFM as it eliminates friction and drag between the sample and tip. Again a constant oscillation amplitude is maintained and the cantilever motion is stored to obtain the topography. The majority of AFM imaging in this work was conducted in tapping mode.

3.4.2 Image Artifacts

With all microscopy techniques the user needs to be aware of the associated artifact mechanisms. For AFM imaging, the resultant image is a convolution of the AFM tip and the imaged object's topography. When the radius of the tip is small with respect to the dimensions of the sample, the effect of the finite tip radius is negligible; however, as sample dimensions approach those of the tip, the image begins to deviate significantly from the original object.

Tip shape is the major factor limiting resolution for AFM. The tip radius, the tip sidewall angles and the tip length contribute to the quality of the image. The smaller the tip radius the higher the resolution. As AFM relies on the interaction of atomic forces between the tip and sample, reducing the lateral interaction area improves the lateral, and vertical resolution. Ideally the tip would act as an infinitely sharp delta function to obtain the best image; in reality the tip radius is typically between 5–10 nm. During scanning resolution can degrade due to tip wear, tip breakage, or the accumulation of contaminants; these increase the effective tip radius, which in turn broadens the image.

The tip sidewall angles become important when imaging objects with steep sidewalls. The AFM is unable to measure object profiles that are steeper than the sidewall angle of the tip. Silicon tips are most commonly used in tapping mode, and as these are etched, the tips have angles of 17° in one scanning plane, and 25° and 10° in the other, according to the silicon etch planes [73]. The silicon tip sidewall angles are illustrated in Fig. 3.8(a) and (b), and the resultant scan of a tip on a sample with vertical sidewalls is illustrated in Fig. 3.8(c). The tip is unable to track along the profile, resulting in an image with sidewall angles the same as the tip.

The tip length can also be restricting for objects that have high surface relief, and small lateral dimensions. Special tips that are manufactured by focused ion beam techniques are available specifically for these applications, however the sidewall angle is still a limitation. Using carbon nanotubes as tips is being researched as they have the desirable properties of a small radius, and steep sidewalls [75].

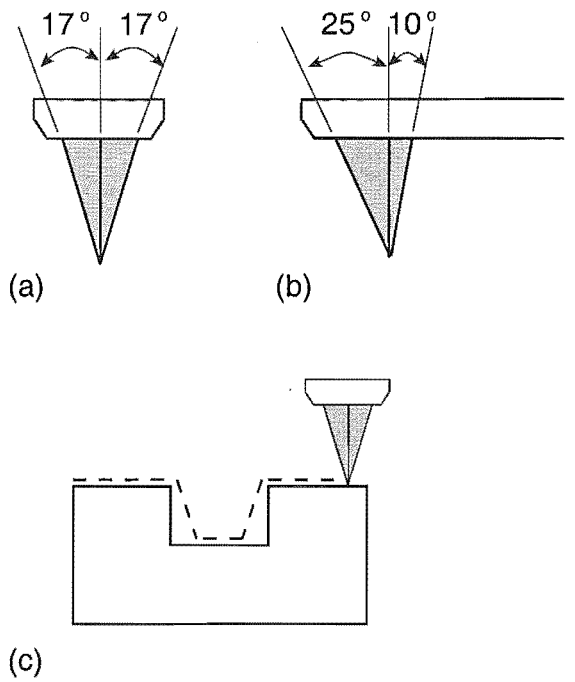


Figure 3.8 Schematic of silicon AFM tips, (a) and (b) show two views of the tip, illustrating the tip sidewall angles corresponding to scan directions of 0° and 90°. (c) illustrates the scan path of a tip along a sample with vertical sidewalls; the scan direction is consistent with the view corresponding to (a).

3.5 SUMMARY

The main techniques used for the experimental work reported in the following two chapters have been described here. These include techniques for the manufacture of ENFOL masks and pattern transfer following ENFOL exposures, and importantly for imaging results at each stage. The pieces of equipment described here are, in the main, recent acquisitions to the group, arriving within the time frame of this thesis work. This meant some effort was involved in establishing basic protocols and in understanding the capabilities of the instruments. This occurred in parallel with the ENFOL research.

Chapter 4

ENFOL MASK FABRICATION

4.1 INTRODUCTION

The membrane mask plays a critical role in the evanescent exposure. The membrane must be conformable so the pattern on the membrane can be in intimate contact over large areas with the substrate. To test the resolution capabilities of ENFOL, it was essential to pattern masks with very small features. This is not trivial, as the patterns must necessarily be much smaller than what our optical lithography equipment is capable of. Electron beam lithography is known for its high resolution capabilities and using this technology patterns down to 70 nm feature sizes were defined onto membrane masks.

This chapter describes the fabrication of the membrane masks. The three main steps are membrane formation via a KOH wet-etch, pattern definition via electron beam lithography and pattern transfer via lift-off metallisation. These are summarised in Fig. 4.1. Mask characteristics such as their transmission, pattern profiles and durability are also discussed.

4.2 MEMBRANE FORMATION

Silicon nitride coated silicon wafers were purchased from the University of California, Berkeley. The silicon nitride is deposited onto both sides of the silicon wafer using Low Pressure Chemical Vapour Deposition (LPCVD) using N_2, SiH_2Cl_2, NH_3 gases at a process temperature of 835°C. The deposition is controlled to ensure a silicon rich composition is formed to produce a low stress non-stoichiometric Si_xN_y layer so that even thin layers are relatively robust. Stress measurement data from Berkeley indicate the Si_xN_y layer has a residual stress between 300 and 355 MPa. Si_xN_y thicknesses of both 1 μm and 2 μm on the front side were trialed. The thicker membrane was found to be more robust, consistent with what is reported in the literature [76]. Initially the wafers were purchased with the backside Si_xN_y removed, however on later wafers a 200 nm thick layer of Si_xN_y on the backside of the wafer was left to act as an etch mask during membrane formation.

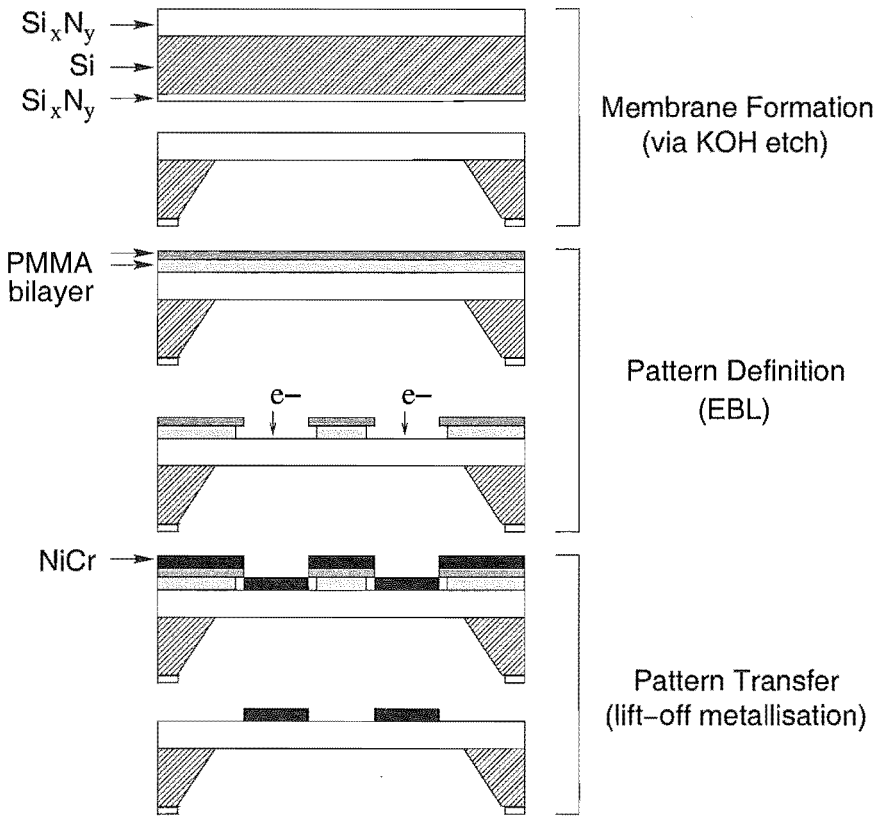


Figure 4.1 ENFOL membrane mask fabrication.

The following steps are followed to form membranes prior to patterning:

Step 1: Coat frontside Si_xN_y *in photoresist.*
Coating the frontside Si_xN_y protects the wafer from contamination during scribing. Typically Shipley S1805 or S1813 photoresist is spun on in the photoresist spinner followed by a short bake of approximately 10 to 15 minutes. Longer bakes tend to make the resist difficult to remove.

Step 2: Scribe the wafer into 10 mm by 10 mm square samples.
A wafer scriber with a diamond tip is used to break the wafer up into 10×10 mm^2 samples. Scribe lines should be made parallel to the wafer flat which is aligned with the silicon crystal planes. This results in more cleanly cleaved samples.

Step 3: Defining the aperture in the backside Si_xN_y.
An aperture is defined by evaporating a NiCr layer of approximately 100 nm, with a circular shadow mask in the centre of each substrate. In this case metal discs with a diameter of 4 mm were used.

Step 4: Dry etch the exposed Si_xN_y.
RIE is used to etch the exposed backside Si_xN_y, with the NiCr acting as an etch mask.

The etch parameters are given in Table 4.1. While an etch rate of 110 nm/min is quoted, a 10 minute etch ensured good clearance of the backside Si_xN_y for etching a number of samples at once.

Table 4.1 Dry etch conditions for removal of backside Si_xN_y.

Gases	Etch Rate (nm/min)	Pressure (mTorr)	Flowrate (sccm)	Power Density (W/cm^2)	Temperature (°C)
CHF_3/O_2	110	55	40/15	0.62	22

Step 5: Wet silicon etch.
A 30% by weight solution of KOH in de-ionised H_2O is used to etch the exposed silicon with the silicon nitride acting as an etch mask. The etch takes 6 hours to remove the approximately 500 μm thick silicon layer. The first 4 hours are etched at 87°C followed by 2 hours at 80°C. This is an anisotropic etch and the silicon is etched from a circular hole to a rectangular window by the preferential etching of the (100) silicon wafer. Figure 4.2 is a photograph showing what a typical membrane mask looks like. It has been patterned using electron beam lithography, and the metallised features are faintly visible.

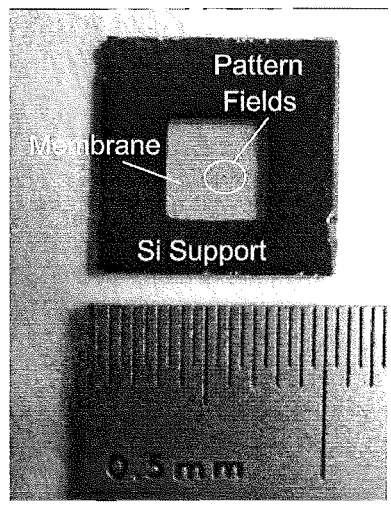


Figure 4.2 Patterned ENFOL membrane mask showing membrane region and Si supported region.

4.2.1 Membrane Transmission

The transmission properties of the Si_xN_y membranes were measured in the Cary spectrophotometer, as outlined in Section 3.3. Table 4.2 gives the transmission through a 2 μm membrane for the relevant spectral lines of a mercury arc lamp. A typical spectrum for a mercury arc lamp is plotted in Fig. 4.3.

Table 4.2 Transmission intensity through a 2 μm membrane measured at wavelengths equating to the spectral lines of a mercury arc lamp.

λ (nm)	Spectral Line	Transmission %
436	G-line	60
405	H-line	29
365	I-line	9
$\leq$ 334		<2

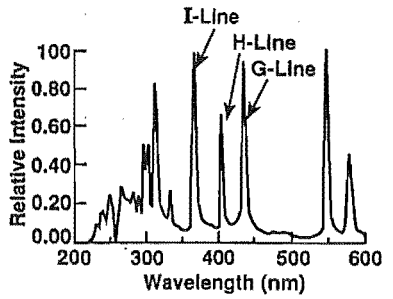


Figure 4.3 Typical high pressure Hg-arc lamp spectrum. Reproduced from [56], p. 33.

The combination of the Hg spectral lines and the low pass filtering nature of the Si_xN_y show that the G-line at 436 nm is the most significant exposing wavelength in the blue-UV range, with significant contributions from the H-line at 405 nm and a decreasing contribution from the I-line at 365 nm. As is evident from the table, the Si_xN_y becomes strongly absorbing below 400 nm making it an unsuitable material for deep UV exposures. These measurements are important as they highlight which wavelengths contribute to the ENFOL exposure.

4.3 PATTERN DEFINITION

The Philips electron beam lithography (EBL) tool, described in Section 3.1, was used to define the mask patterns. EBL is the dominant technology for mask making both commercially and in research. It fits this niche well as it has the advantage of high resolution and its major disadvantage, that of throughput, is not as critical for low volume mask making applications that have a high level of re-use.

4.3.1 Electron Beam Resist Preparation

Membranes are prepared for EBL by spinning a bilayer of PMMA, a high resolution organic electron sensitive resist. The first layer consists of either a 4% or 3% low molecular weight (LMW) PMMA with an average molecular weight M_w = 120k dissolved

in chlorobenzene solvent. This is then spun on the membrane at 3000 rpm followed by a bake in a convection oven at 185°C for 30 minutes. A second layer consists of 2.5% high molecular weight (HMW) PMMA with an average $M_w = 996$k dissolved in xylene solvent. This is then spun on at 3000 rpm, followed by a further 30 minute bake at 185°C.

A modified chuck was required to spin the PMMA onto the membranes, to prevent direct vacuum being applied to the membrane windows. A plate with holes for vacuum corresponding to the outer silicon support area satisfied this criteria. During the spin cycle there is some deformation of the membrane that results in an uneven distribution of resist. The non-uniformity is clearly visible as colour variations that are particularly evident at the membrane corners and edges. However the resist in the central area of the window remained uniform so variations in exposure results could be avoided by restricting the EBL exposure to this central area.

The bakes drive off the solvents in the PMMA solution. They must be at a temperature above the glass transition temperature (T_g) of PMMA (105°C), and above the boiling point of the casting solvent, but below the decomposition temperature of PMMA (200°C). The glass transition is a second-order transition that marks the solid changing from a brittle glass-like material to a flexible plastic [77].

During the exposure process, electrons cause chain scission of the PMMA, reducing the molecular weight of the polymeric chains. The exposed areas are then more soluble and are dissolved more quickly than unexposed areas during development. The PMMA bilayer is used to improve the profile of the resist following exposure. The LMW PMMA layer dissolves more quickly during development than the HMW PMMA layer, creating an undercut profile appropriate for lift-off. An undercut profile reduces the amount of metal that gets deposited on the sidewalls by providing a degree of shadowing and reducing problems with tearing during lift-off.

Originally both the LMW and HMW were cast in a xylene solvent. However, for LMW percentages of 4% and higher, xylene did not dissolve the PMMA adequately and a stronger solvent, chlorobenzene, was required. Having the top PMMA layer cast in a weaker solvent is desirable as it reduces the likelihood of the bottom PMMA layer being partially dissolved during its application [78].

4.3.2 Exposure and Development

A series of pattern files were generated, by the process described in Section 3.1.1, for feature sizes ranging from 1 μm down to 70 nm. A typical pattern file for one particular feature size consists of 50% duty cycle gratings, sparse 25% duty cycle gratings and dense 75% duty cycle gratings. For example the 250 nm feature size field is shown in shown in Fig. 4.4. On the right is a region of 250 nm features on a 500 nm period, in the centre a dense field with 750 nm features on a 1 μm period, and on the left a sparse

field with 250 nm features on a 1 μm period. These patterns were designed to test the ultimate resolution of ENFOL for grating structures as well as to investigate the effect of duty cycle.

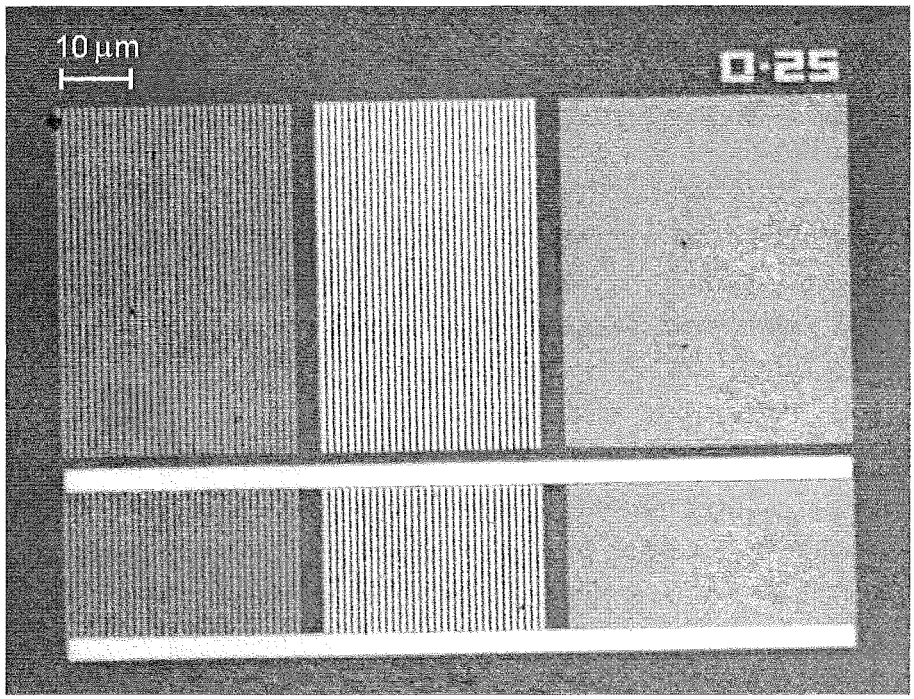


Figure 4.4 Optical micrograph of a membrane mask pattern for 250 nm feature sizes. On the left are sparse gratings with 250 nm metal lines and 750 nm apertures, in the centre are dense gratings with 750 nm lines and 250 nm apertures and on the right, 250 nm lines and apertures.

Features were predominately written with $\times 1250$, $\times 2500$, or $\times 5000$ magnifications, with corresponding field sizes ranging from 115 μm by 75 μm, down to 30 μm by 20 μm. This range of magnifications was chosen as the pixel sizes were smaller than the physical beam spot size, but not unreasonably so. Working at higher magnifications, equating to smaller pixel sizes is of diminishing benefit when the beam size is significantly bigger than the pixel size. Field sizes are also reduced at high magnifications. Working at lower magnifications can result in pixel sizes greater than the beam, which suffer from underexposure at pixel edges. Typical exposure parameters are given in Table 4.3.

Table 4.3 Electron beam lithography parameters.

Voltage	Beam Current	Spot Size (Setting)	Spot Size (Actual)
50 kV	50 pA	160 Å	400-500 Å

Following exposure the samples are developed in a solution of 3:1 isopropyl alcohol

(IPA):methyl-isobuty-ketone (MIBK) at 23°C for 30 seconds. To halt development they are then rinsed in pure IPA. During the development process the lower molecular weight chains (the area exposed to electrons) are dissolved away.

4.3.3 Dose Optimisation

The best exposure dose for a pattern is dependent on a large number of parameters: the substrate's atomic mass and conductivity, magnification, beam size, pattern density and layout. This assumes operation at the optimal focus with minimal astigmatism. Drifts in the beam current over time can also affect results. To optimise the exposure, identical fields with varying exposure doses were evaluated, with all other parameters as fixed as possible.

Exposures onto membranes require higher exposure doses (between 10-20% higher) in comparison to thicker silicon or silicon nitride coated silicon samples due to a reduction in scattering from the thin membranes. This is particularly evident at small feature sizes, which need larger exposure doses to get good resist clearance during development.

Imaging membrane masks is difficult which made optimisation of the exposure dose awkward. The optical microscope is a useful tool to some extent for evaluating the written pattern, but as the smallest patterned features were smaller than the resolution limit of the microscope, it was not always clear if the patterns had been exposed and developed adequately. With experience it is possible to make a good guess however. Ideally, the exposed and developed pattern in the PMMA would be imaged to evaluate the exposure, but this was a problem for a number of reasons. Initially the only imaging tool with high enough resolution to evaluate the PMMA patterns was an SEM. However, as the membranes are insulating, they charge-up making imaging impossible and sometimes breaking the membranes. Coating the PMMA after development in a thin Au layer helps to alleviate the charging, however organic materials such as PMMA are not stable during the imaging and tend to heat up and warp. This is particularly noticeable for small features. Encapsulating the resist in a plasma enhanced chemical vapour deposited film at room temperature has been reported to stabilise the resist for imaging [78], however facilities to do this were unavailable.

Evaluating the patterns following lift-off indicated to some extent the best exposure dose. An underexposed field (no metal remaining) was easy to recognise, but overexposed areas were difficult to discern. The availability of an AFM in the latter stages of this research enabled imaging of the PMMA down to the smallest feature-sizes, in a non-destructive way and with minimal sample preparation. This enabled dose optimisation trials to be conducted more successfully although awareness of AFM imaging artifacts was essential. The scanned images are strongly reliant on the quality of the tip which can wear, get contaminated or break. Re-imaging the same field with a

new tip can yield quite different results; fields that initially appeared to be unresolved would miraculously appear. The measured depth of the PMMA patterns can change noticeably depending on the physical size of one tip compared to another as well as the magnitude of the force applied for small features. Experience of AFM imaging helps ensure the images are correctly interpreted. This is particularly important when the tip size is significant with respect to the sample patterns, which is the case for the 100 nm scale features and below, with relatively high aspect ratios.

The smaller period gratings, 130/260nm, 100/200nm, 70/140nm lines/period were the most difficult patterns to write with the EBL, and the most sensitive to exposure dose. The best resolutions were achieved by writing gratings with a smaller than 50% duty cycle, with a reasonably high exposure dose. This was partly due to the larger beam diameter compared to the pixel size. An estimate of the number of pixels to write per grating period was calculated from

$$no.\ pixels = \frac{l - (d_b - pix)}{pix}, \tag{4.1}$$

where l is the linewidth (half the period), d_b is the physical beam diameter, and pix is the size of a pixel in nanometres. This equation takes into account the effect of the beam diameter being broader than the pixel size. Table 4.4 lists the duty cycles and exposure doses that were found to give good results at magnifications of $\times 2500$ and $\times 5000$. The written duty cycle decreases with the feature size, as the beam diameter becomes more significant relative to the feature size.

Table 4.4 Exposure dose versus line size at a magnification of $\times$ 2500 and $\times$ 5000.

Magnification	Field Size (nm)	Line Size (nm)	Written Duty Cycle (%)	Exposure Dose (μC/cm^2)
2500	58 × 39	170	42	800 – 900
		130	37	1000
		100	29	1200 – 1800
		70	20	1800 – 2000
5000	28 × 19	100	28	1200 – 1600
		70	20	1600 – 1800

Figure 4.5 illustrates an AFM image of nominally 100 nm lines on a 200 nm period grating in 155 nm thick PMMA. The PMMA was exposed with a dose of 1800 μ C/cm^2 at $\times$ 2500 magnification. The image was taken in tapping mode with a silicon tip. The flat bottoms of the trenches indicate that the tip is getting right to the bottom of the trench, and the depth of the trenches suggest the resist is fully cleared. The profile does not appear to be undercut or even vertical but this is an artifact of the AFM, where the steepest slope that the AFM can measure is the slope of the AFM tip. While an undercut profile hasn't been imaged, subsequent lift-off in well-exposed fields

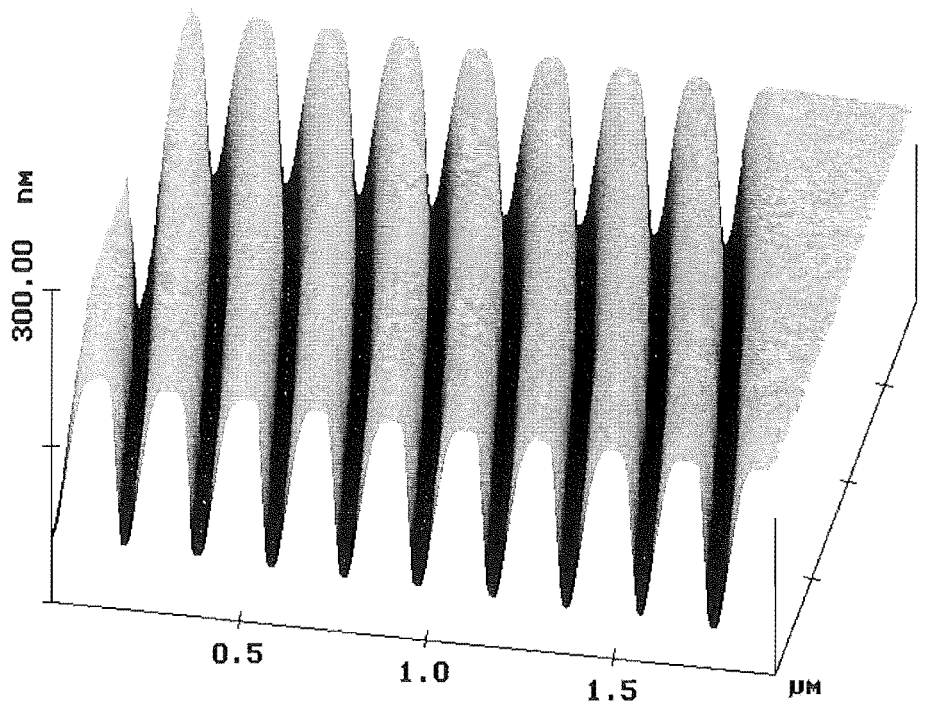


Figure 4.5 AFM micrograph of a 200 nm period grating in PMMA.

is generally good indicating that a good resist profile has been obtained. Figure 4.6 shows a 140 nm period grating exposed with a dose of 3000 μC/cm^2 at $\times$5000 magnification. The bottoms of the trenches are quite pointed and it is difficult to know if the tip is reaching the substrate. Some of the tops of the PMMA have been developed away by approximately 10 nm compared to the background PMMA, a small section of which is displayed in the right-hand side of the scan. This indicates there was some overexposure, however the deepest part of the trench is just over 150 nm, suggesting full development or very close to it. Using a thinner PMMA bilayer may make it easier to expose the 140 nm period grating without thinning of the top PMMA layer, however lift-off is then more difficult.

4.4 PATTERN TRANSFER

Pattern transfer of the PMMA image into a 30–40 nm thick NiCr layer was performed using a lift-off process. A metal film is evaporated onto the sample and then a solvent is used to remove the PMMA and all metal not adhering directly to the Si_xN_y membrane.

Firstly the sample was given a 10 second O_2 plasma ash to remove any residual organic material in the cleared areas to enhance adhesion. The ash was performed at 500 W, 1 cc/min flowrate and a pressure of 0.3 Torr to give an etch rate of $\approx$ 1 nm/s.

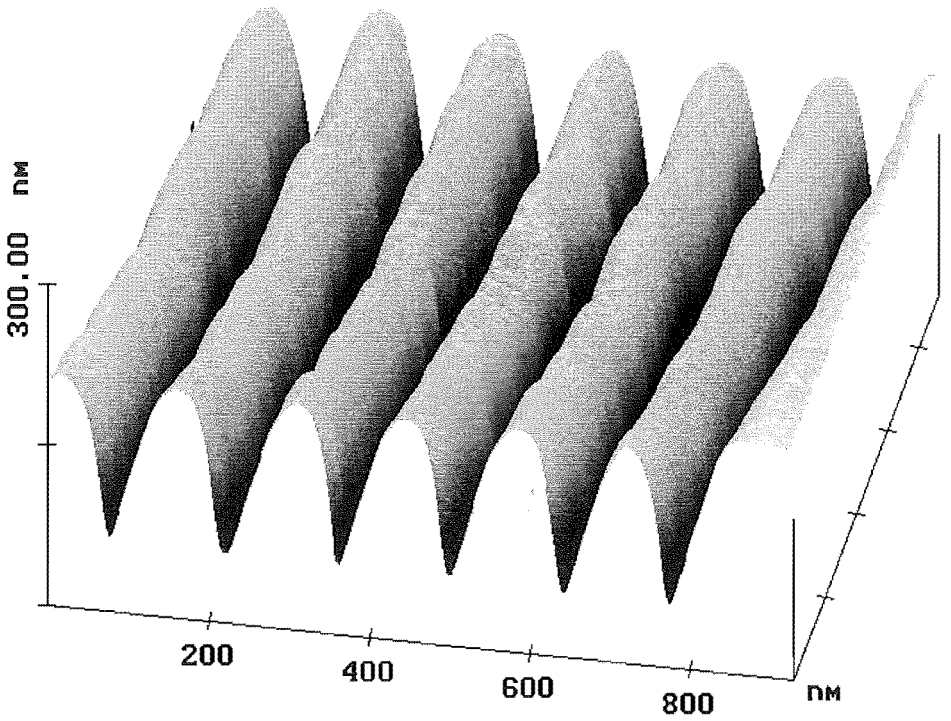


Figure 4.6 AFM micrograph of a 140 nm period grating in PMMA.

NiCr with a ratio of (Ni80/Cr20) was used, as it has a small grain size and good lift-off characteristics down to 100 nm-scale feature sizes. Some metals such as Al do not lift-off as easily for small features and Au has a large grain size which can create additional edge roughness for small features. Ideally a good conductor with a small skin depth is preferred for ENFOL. Tables of NiCr physical properties were unavailable, and taking thin film measurements to get accurate results is a difficult process requiring good knowledge of the material properties of the substrate the NiCr is applied to. However, the NiCr metallisation performed adequately for the initial ENFOL studies.

The NiCr is deposited by a physical evaporation process. The metal is heated in a chamber at low pressures to a temperature high enough that it is vapourised. The metal is then redeposited on surfaces placed within range. A quartz crystal in the evaporation chamber monitors the thickness of the metal deposited. The crystal's resonant frequency changes proportionally with the increase in mass of material deposited on it, and the thickness can be calculated from the known density of the material. NiCr changes the resonant frequency by 4 Hz/Å, and by placing the sample at a similar distance from the source as the crystal, the deposited thickness can be deposited in a controlled way.

NiCr wire is cut into small pieces and placed in an alumina coated molybdenum

boat. The membrane sample is clipped to a chuck and placed as directly as possible above the evaporation boat. This ensures that when the metal is vapourised, it is incident at an angle as close to normal as possible. Angled evaporation is to be avoided, as it results in poor lift-off as one side of the sidewalls gets coated in metal making the lift-off stage more difficult.

The evaporation chamber is pumped down in a two stage process. Firstly the chamber is heated to encourage more efficient out-gassing. The chamber is then cooled to further reduce the pressure, and to condense out any remaining contaminants. A low pressure is required to prevent scattering of the evaporant as well as to reduce the presence of contaminants that may be unwittingly deposited. Once a pressure of 10^{-6} Torr is reached, power can then be applied to heat the boat to temperatures that will vapourise the metal. The power is increased slowly to avoid thermally shocking the boat. Boats are generally made of a refractory metal that can be heated to high temperatures without being melted and will be unreactive with the material to be evaporated. The boat's alumina coating improves the deposition as it encourages the NiCr to contract into a tight ball effectively decreasing the size of the evaporation source. The smaller the source, the narrower are the range of angles at which the evaporant arrives at the sample, thereby improving the lift-off. Initially a shutter shadows the boat from the sample (and crystal) to avoid deposition from any contaminants on the NiCr as well as thermally shielding the samples. It also provides some control on when to begin the deposition, as evaporating at different pressures can result in differences in the composition of the deposited material. Once a suitable power level is reached (which changes with the boat size, age and condition), the shutter is opened until the desired thickness has been deposited.

The final step in the lift-off process is to leave the sample in acetone to dissolve the remaining PMMA. Initially this step was performed at approximately 60°C but since then improved results have been achieved at room temperature. While heating the acetone speeded up the lift-off process, the NiCr crazed and came off in small pieces some of which redeposited on the membrane. At room temperature it lifts off in an almost continuous layer.

Figure 4.7 and Fig. 4.8 show AFM images of a 340 nm period grating and a 200 nm period grating respectively in 30 nm of NiCr. The images show well defined lines with little line edge roughness. The top surfaces of the metal gratings are non-uniformly rough however and may adversely affect the ENFOL exposure.

4.5 MASK PATTERN PROFILE

The mask patterns' profile is a potentially important characteristic for a conformable contact exposure technique such as ENFOL, where the substrate-mask distance is critical. The unwanted topography in the mask patterns' profiles was only observed fol-

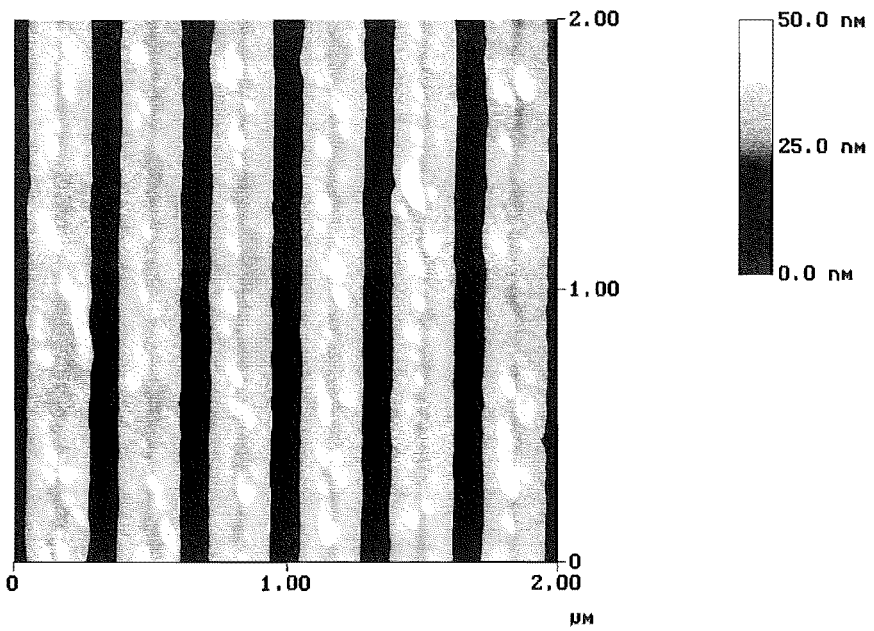


Figure 4.7 AFM micrograph of a 340 nm period metallised grating.

lowing the acquisition of the AFM, quite well into this research. Previous SEM images showed no indication of non-uniform topography on the mask patterns. Figure 4.9 illustrates a 200 nm period metallised grating on a membrane imaged in an SEM after it broke[1]. The grating edges look smooth and regular, and there is no indication of any topography on the metal surface. The slightly shadowed appearance of the image is a result of minor charging.

In Fig. 4.7 the metal appears to have curled back on itself, raising the height of the metal. This topography is undesirable as it will increase the separation between the mask and substrate during exposure. It may not be a problem from a diffraction perspective if it is uniform, however the height of the extensions varies typically from 10 - 20 nm.

Figure 4.8 illustrates a different problem with the mask topography; here the metal has been successfully lifted off but has been redeposited onto the remaining patterned metal. This profile is probably worse than the previous profile, as during an exposure the redeposited metal will make it difficult for the substrate to be in conformable contact with the mask pattern, introducing a small air-gap at the pattern edges.

Clearly, the more successful the lift-off, the better the mask pattern profile. Two factors determine the success of the lift-off, the angle at which the evaporant hits the mask, and the profile of the bilayer PMMA. The mask substrate was aligned by eye

[1]The membrane edges were coated in a colloidal graphite paste, and the membrane was stuck down on carbon tape to reduce charging.

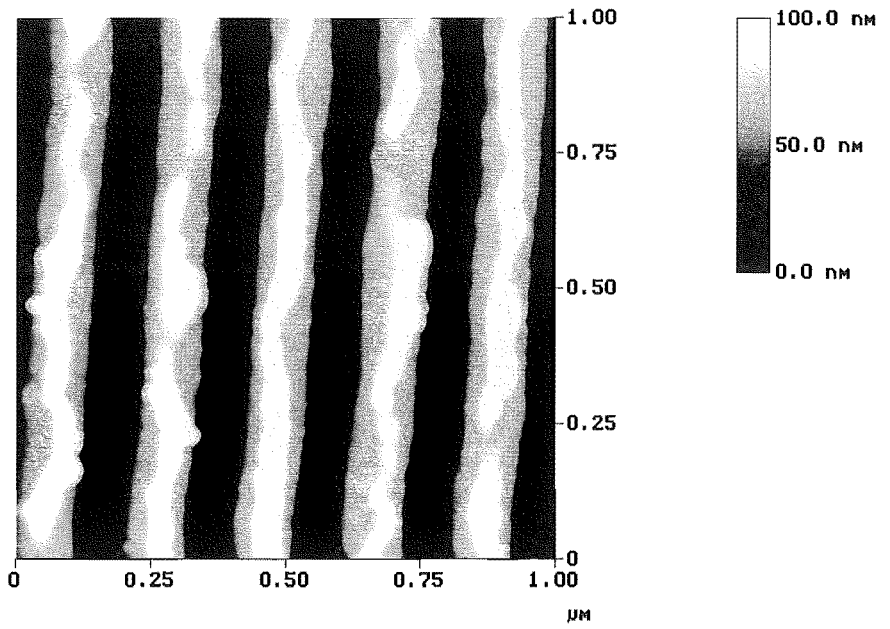


Figure 4.8 AFM micrograph of a 200 nm period metallised grating.

to be as directly above the source as possible. Small variations in the incident angle of the evaporant during metal deposition may have contributed to lift-off problems, however as the profile was generally symmetrical, it is more likely to be due to resist profile. Some heating during the deposition could be responsible for some warping of the PMMA that may have degraded the profile somewhat, given that the T_g is around 105°C, and high temperatures are required for melting the metal. An unstable plasma during the ash to remove organic residue may also have resulted in inconsistencies in the thickness of PMMA etched. While it is desirable to remove the organic residue, over-etching can reduce the bilayer resist, degrading the undercut profile.

Further investigation into improving the resist profile is necessary for consistently good lift-off. Tighter control of the plasma etch (in the RIE perhaps) and perhaps a more heat isolated deposition system such as electron beam deposition may improve the profile. Changing to a subtractive process to define the absorber patterns may be a solution to mask profile issues. In general to etch metals, chlorine gas is required which is unavailable with the current dry etch facilities. Tungsten may be a possible absorber alternative however, which is more readily etchable.

4.6 MASK ROBUSTNESS

Membrane mask robustness turned out to be one of the foremost issues for ENFOL. Frequent mask breakages inhibited a systematic experimental characterisation of the ENFOL process. While membrane masks can be produced with very few breakages, the

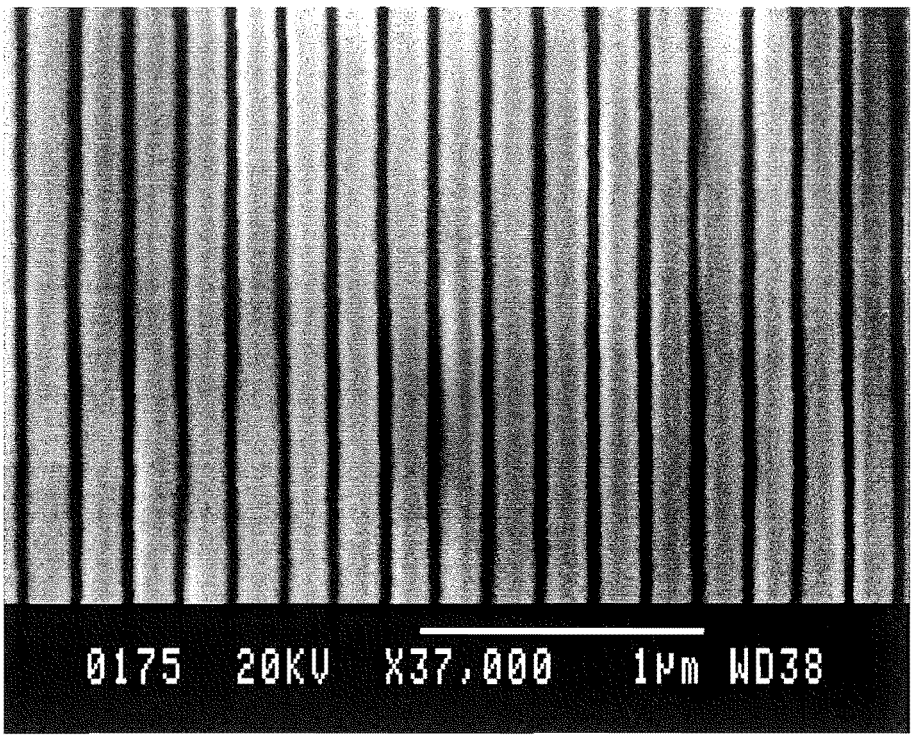


Figure 4.9 SEM image of a 200nm period metallised grating. The membrane was imaged after being broken. It was coated by colloidal graphite at the edges and stuck to a carbon tape for imaging.

ENFOL exposure process has a high attrition rate. Mask lifetimes are variable; a mask might break on its first exposure, or as in the case of my most hardy mask, survive for as many as twenty exposures. This suggests that improving mask robustness should be possible, and some work has been carried out to this end.

Commonly mask breakages occur at the membrane window edge. This is thought to be because in this region, the membrane is least able to conform over topography. Here the silicon makes a $\sim 54.7°$ angle with the Si_xN_y due to the anisotropic silicon etch in KOH. The abrupt transition from a flexible window to a solid Si_xN_y on silicon, produces a small bend radius that concentrates stress. Jagged edges can be created at the membrane/support boundary, the sharp points being potential sources of membrane weakness. The jaggedness occurs during the anisotropic etch and results from the membrane boundary forming along different Si crystal planes. This is illustrated in Fig. 4.10.

Trimble *et al.* reported that by using an isotropic etch to create a shallow skirt in the membrane support, fracture pressure differentials were improved by a factor of six [76]. Creating a shallow isotropic profile makes a more gradual transition from the flexible membrane area to the rigid support giving a larger flexing radius.

Some initial experiments were performed to see if the mask robustness could be improved by incorporating an isotropic etch to taper the silicon support. To form the

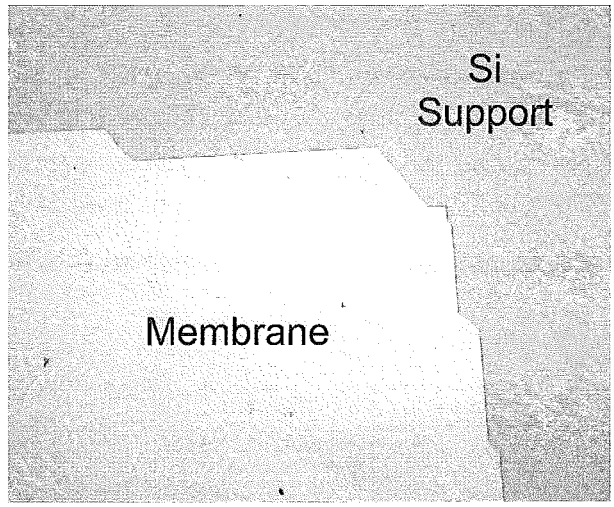


Figure 4.10 Jagged boundary of between Si_xN_y membrane and silicon supported region. The jagged outline results from misalignment of the boundary and Si etch planes during the anisotropic Si etch.

silicon nitride membranes, a thickness of approximately 500 μm of silicon needs to be removed from the central window region. Firstly the standard anisotropic KOH etch was used to etch the bulk of the silicon, as this had a high etch rate and good selectivity between the silicon and silicon nitride. When $\sim 100\,\mu$m of the silicon remained, an isotropic etch was used to clear the membrane and obtain a gentle skirt. A number of different etchants for etching silicon isotropically have been reported in the literature and a summary is provided in Table 4.5. Of the isotropic etchants trialed, the HNO_3:HF:$NaNO_2$ gave the best results. The HNO_3:HF:H_2O etchant's etch rate was too slow and HNO_3:HF was rejected outright due to the high HF concentration that would etch the Si_xN_y.

Table 4.5 Isotropic silicon etches with reported etch rates for bulk silicon.

Etchant	Ratio	Reported Etch Rate	Reference
HNO_3:HF:$NaNO_2$	95 ml:5 ml:1 g	4 μm/min	[79]
HNO_3:HF:H_2O	150 ml:8 ml:75 ml	-	[80]
HNO_3:HF	21 ml:4 ml	13.8 μm/min	[81]

The silicon etch rate for the chosen HNO_3:HF:$NaNO_2$ etch was $\approx 1.6\,\mu$m/min, lower than the quoted $4\,\mu$m/min. This was achieved by heating the solution to around 70°C and stirring. The etched surface should be facing up, (membrane surface down), to allow bubbles to be released to obtain the best etch rate. During the etch, some of the Si_xN_y is removed, however the etch rate is less than 200nm/hour. Note that during etching, a holder that promotes even etching over the front and back surfaces of the Si_xN_y is important. Unevenness in the etching of the top or bottom surfaces must be avoided as during the spin coating of the membranes the PMMA will pool unevenly,

resulting in pattern irregularities.

Figure 4.11 illustrates the difference between the profile of the silicon supports for isotropically etched Si (Fig. 4.11(a)) and anisotropically etched Si (Fig. 4.11(b)). Both samples have been cleaved, and the SEM view is of the cross-section, with the silicon support on the right. In Fig. 4.11(a) the silicon has not been etched through to form

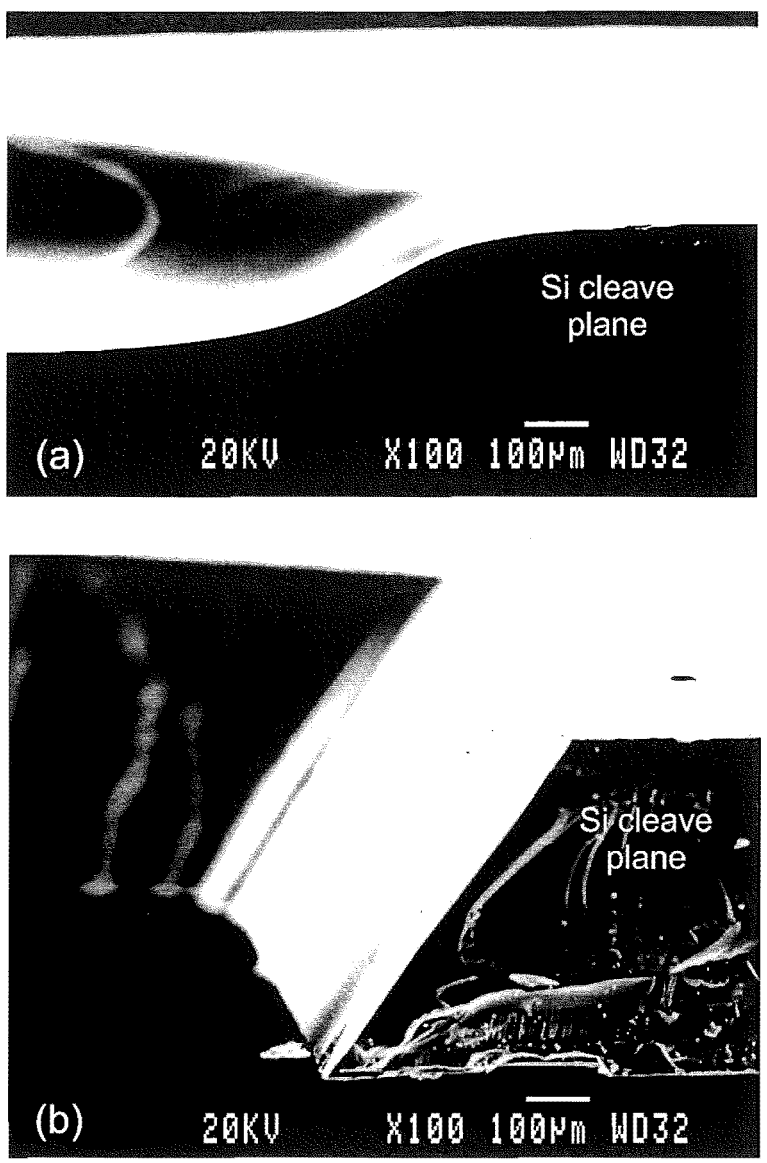


Figure 4.11 SEM Cross-section of membrane edges for (a) isotropic etch in HNO_3:HF:$NaNO_2$, and (b) anisotropic KOH etch.

a membrane. It was etched without the anisotropic pre-etch, and the backside Si_xN_y which had a circular window defined in it, has been totally etched away. The top bright surface is therefore the back of the Si, and a circular depression can be seen on the left, from the circular window defined in the backside Si_xN_y. A very smooth tapered profile is evident in the foreground, (dark to light boundary). In Fig. 4.11(b) the silicon has

been etched right through, the originally circular window defined in the backside Si_xN_y has become a square due to the preferential etching of the Si crystal planes. The angle of the silicon sidewall support, (visible in the foreground) is an abrupt $\sim 54.7°$.

Membranes were successfully made using a combination of anisotropic and isotropic etching. One was subsequently patterned and used for ENFOL exposures, without breaking. Due to time constraints, a large enough trial was not conducted to assess the improvement in mask robustness, but initial results appear promising.

4.7 SUMMARY

A fabrication process for developing membrane masks has been described. The masks fulfil the requirement of being conformable over surface topography, and mask patterns have been defined with dimensions small enough to test the resolution of ENFOL. Feature sizes down to 70 nm have been patterned, close to the limit of our EBL resolution. While a process for making membrane masks has been successfully developed, some deficiencies have been highlighted along the way.

AFM imaging indicates that the profiles of the mask patterns are not ideal. The non-flat profiles may increase the mask–substrate separation and may contribute to line edge roughness. To improve the profile, tighter control of the resist profile and lift-off will be required. Some tearing is difficult to avoid, particularly when patterning sub-100 nm structures. An alternative would be a subtractive pattern transfer stage for defining mask patterns. Tungsten could be a suitable material for this.

The lack of robustness of the membrane masks is perhaps the most serious problem that was discovered. While they can now be fabricated with high yield, they suffer from an unacceptably high attrition rate during the contact exposure. Some efforts have been made into improving the robustness of the masks by tapering the silicon supports surrounding the membrane. In the future if larger area masks are to be made, some changes to the masks will be required. Thicker membrane materials may be necessary, as the fracturing force is proportional to the square of the membrane radius. Thin silica masks may be an alternative mask substrate. Conformable contact lithography has been reported using silica $\sim$150 μm thick; the increased thickness of the silica compared to the membrane masks making them more robust [46].

While there are refinements to be investigated and incorporated in the mask making process, masks have been successfully fabricated and used for investigating the concept of sub-diffraction limited contact exposures. These experiments are discussed in the following chapter.

Chapter 5

ENFOL EXPOSURE AND PATTERN TRANSFER EXPERIMENTS

5.1 INTRODUCTION

This chapter describes experiments that were conducted to test the resolution capabilities of ENFOL. In order to obtain high resolution with this technique, exposure must be made into a thin resist layer, as the contrast for small features has a limited range. This thin resist is an essential requirement for ENFOL, but complicates the pattern transfer stage from the photoresist image to more useful materials for device manufacture. For ENFOL to be a viable lithography option, both subtractive and additive pattern transfer techniques must be able to be demonstrated to provide a full processing tool-set.

This chapter discusses the characteristics of ENFOL exposures, investigating the effect of feature size on parameters such as resist thickness and exposure time. In associated work a subtractive pattern transfer process was developed that was compatible with ENFOL exposures into thin resist [52, 53] and for completeness this is briefly summarised. The development of an additive pattern transfer process suitable for thin resists is described. The additive pattern transfer method used is a lift-off metallisation scheme, utilising a trilayer resist system.

5.2 EXPOSURE AND DEVELOPMENT

5.2.1 Experimental Setup

ENFOL exposure trials were conducted using the setup shown in Fig. 5.1. The exposure source was a 200 W high pressure mercury lamp installed in a Cobilt CA-800 mask aligner. This has a broad-band spectrum, with radiation from 300–600nm, as discussed in Section 4.2.1. The Cobilt has a very simple catoptric setup which is illustrated in Fig. 5.2. The resultant illumination is a large-area divergent beam, with a divergence angle of between 3 to 4°. While a collimated source would have been preferred, this was not available, and the small divergence angle was not considered significant.

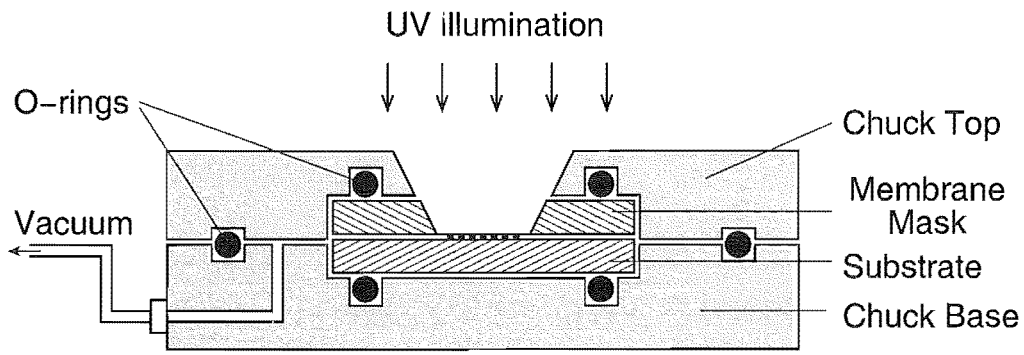


Figure 5.1 Schematic of ENFOL exposure setup.

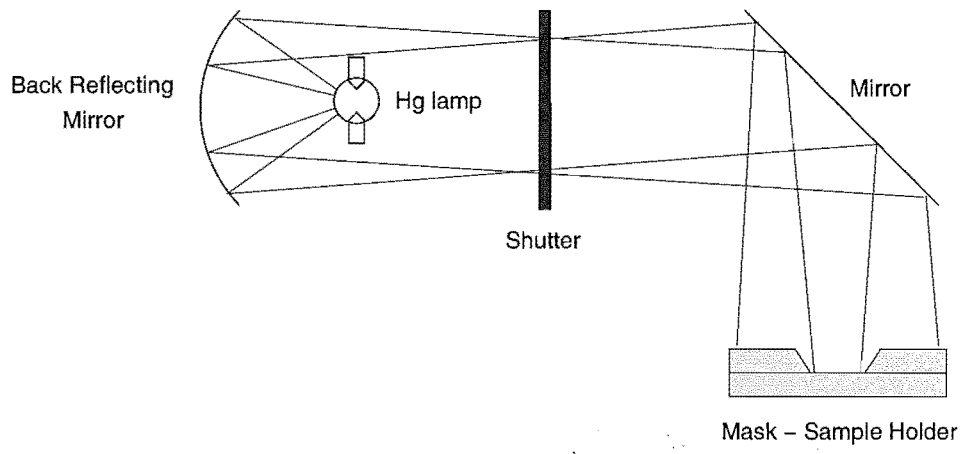


Figure 5.2 Cobilt illumination and optical arrangement.

Very few of the mask aligner's features were used apart from the source, shutter and timer. A custom built chuck was designed to hold a membrane mask and sample under vacuum pressure. Figure 5.3 illustrates the chuck. It is comprised of two machined aluminium holders designed to mount the sample and mask. The bottom section consists of a recessed square for holding $10 \times 10\,\mathrm{mm}^2$ samples with a nitrile o-ring also partly recessed. The top section also has a recessed section for holding $10 \times 10\,\mathrm{mm}^2$ masks, again with a partially recessed nitrile o-ring (not shown). A vacuum line attaches the bottom section to a rotary pump for evacuating the chuck. The top and bottom o-rings act as seals for the vacuum and also allow for some variation in height of the sample/mask thickness. The central area of the chuck, (7 mm diameter) is machined out to allow light through the mask for the exposure. The sample and mask are loaded, then the top chuck is placed onto the bottom chuck with guide pins to minimise any rotational forces that could stress the membrane mask. Once the sample and mask are seated the vacuum is turned on gently which sucks the membrane onto

the substrate. Intimate contact can be confirmed by observing the absence of Newtons rings, interference rings that are visible when there is a varying gap between the mask and sample. When viewing the membrane through the exposure aperture, the mask appears rough as it conforms over particulates on the sample, also indicating that good contact has been achieved.

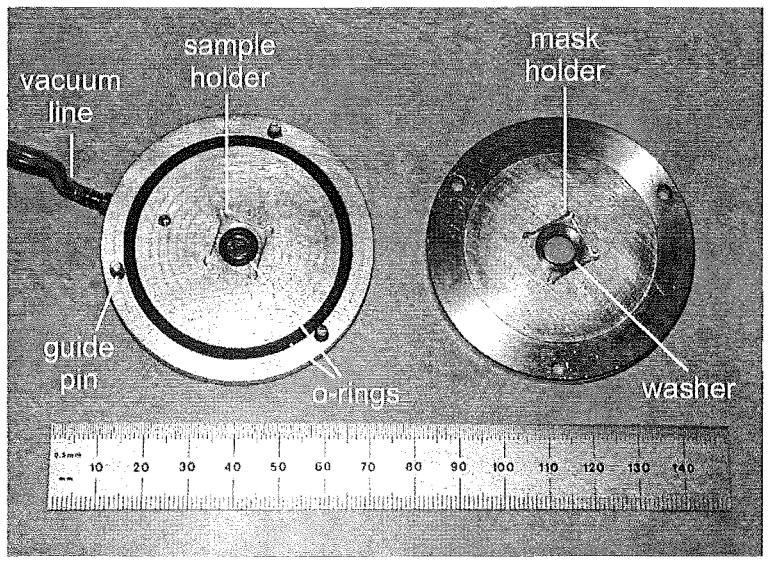


Figure 5.3 Photograph of the ENFOL exposure holder used to keep the sample and mask in conformal contact.

A small range of substrate thicknesses can be catered for by adjusting the thickness of a washer beneath the top o-ring. This allowed both glass substrates, approximately 200 μm thick, and silicon substrates 500 μm thick, to be used. Using a washer of the correct thickness is important. If the washer is not thick enough there is not enough pressure to obtain a good seal for the o-rings, but if it is too thick the excessive force breaks the membrane. Loading the chuck and applying vacuum are the most risky stages in terms of mask breakages. Contamination of the sample and mask needs to be minimised as while the mask is conformable it can only conform to a certain degree; high topography and/or sharp protrusions can puncture the membrane. Contamination or topography near the edge of the membranes close to the silicon supports can be more devastating than in central regions as the conformability is more limited here.

5.2.2 Exposure and Development Parameters

Whilst ENFOL exposures were made into a variety of resist systems, the majority of work employed the same photoresist (PR) layer. This was a commercial G-line resist (Shipley S1805), thinned using a propylene glycol monomethyl ether acetate (PGMEA) solvent. While different photoresist thicknesses were explored, the best results were obtained using a dilution ratio of 1:4 of S1805 to solvent. Silicon samples are spin

coated with the diluted resist at 4000 rpm, followed by a pre-exposure bake at 95°C for 25 minutes. The resultant resist film thickness is around 60 nm. For this thin resist, the optimal exposure time varied between 60 and 70 seconds. Narrow-band UV filters were used to measure the exposure power at 436 nm and at 365 nm. Exposure powers of 93 mJ cm^{-2} and 9.6 mJ cm^{-2} were measured, after the filter attenuation was accounted for. This corresponds to exposure energies of approximately 60 mJ cm^{-2} and 0.8 mJ cm^{-2} at 436 nm and at 365 nm respectively into the resist, once a correction for the membrane transmission of 60% and 9% is incorporated.

A 10 second dip development in Microposit MF320 (an organic developer containing tetra methyl ammonium hydroxide) diluted 2:1 with de-ionised (DI) water was used for most samples. Exposure and development times from previous work with the single layer resist [53] were used as a starting point and the exposure time was adjusted to optimise the exposure, keeping the development time fixed. Long time-scale fluctuations in the source intensity meant that the optimal exposure time was not fixed, but varied over a range of 60 – 70s. Different resist systems also meant the optimal exposure time varied, as the amount of light reflected from the substrate contributing to the exposure varied.

5.2.3 ENFOL Exposure Characteristics Into Thick Resist

To more closely examine the characteristics of the ENFOL exposure, exposures were performed into undiluted Shipley S1805 photoresist, spun to a thickness of $\approx$ 470 nm. This is thicker than the depth of field for the majority of feature sizes, allowing an assessment of the exposure and development characteristics via the topography of the resist. Partial developments have been used previously for a semi-quantitative evaluation of intensity distributions for phase masks [82]. Various exposure and development conditions were investigated for four samples and these are summarised in Table 5.1. The resultant resist profiles were then examined in the AFM.

Table 5.1 Summary of exposure and development conditions for four samples, with exposures made into 470 nm thick photoresist.

Sample	Exposure Time (s)	Development Time (s)	Mask	BRR (nm)
A	65	10	1	290
B	65	10	2	230
C	40	10	2	100
D	20	6	2	100

Three resist parameters were measured, as defined in Fig. 5.4. The background resist reduction (BRR) was measured for each sample, which is the depth of resist developed away in unpatterned regions compared to a totally unexposed region under

the silicon membrane support. Two other resist depths were measured over the different exposure fields with different grating periods. These were the modulation depth a, which is the difference in height between the exposed and unexposed regions of the grating, and the pattern depth d, which is the difference in height between the maximum height of unexposed regions of the pattern compared to the background exposed resist height. In an ideal lithography system, $d = a =$ BRR, indicating that none of the thickness of the nominally unexposed regions is developed away, and full-development has been achieved in patterned regions. However, for ENFOL this is shown not to be the case.

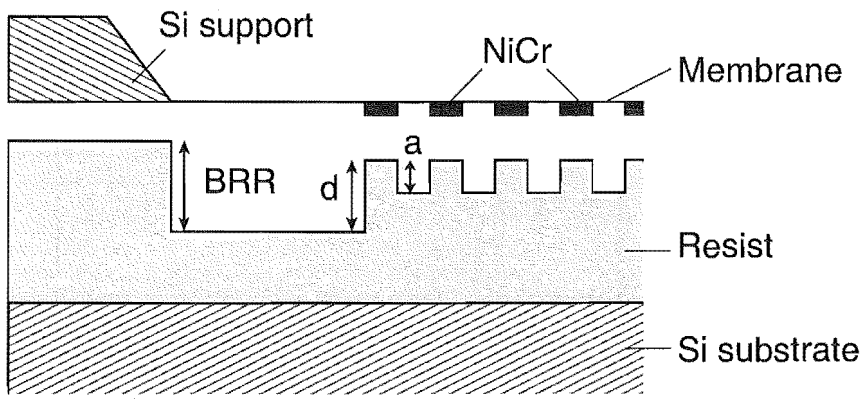


Figure 5.4 Schematic illustrating a resist profile following exposure into thick resist and subsequent development. Measured resist parameters are defined — the background resist reduction (BRR), modulation depth a and the pattern depth d.

Figure 5.5 plots the pattern depth d and modulation depth a versus grating period for the four samples studied. In Fig. 5.5(a) it can be seen that the large period gratings have a d close to that of the BRR. For all samples only the $2\,\mu$m period gratings appear to satisfy the criteria of d equal to BRR. This indicates that there is little decrease in the pattern depth due to transmission through the NiCr mask absorbers. As the grating period reduces, d starts to reduce also as there is increased light leakage around the absorbers. What is interesting about Fig. 5.5(a) are the notable minima that occur at specific grating periods, indicating a higher overall transmission through these gratings. A strong reduction in d is evident in the grating periods from 700 nm to 340 nm for sample A that is well outside the measurement error. The reduction in resist thickness is visible in the optical microscope by a distinct colour change. Sample B shows a similar trend at similar periods, although not as pronounced. The non-linear nature of the resist response to exposure and development conditions may explain the quite large change in depth in this minima region, even though these samples have undergone nominally identical exposure. Samples A and B were exposed using different masks, an additional parameter that may explain some process variability.

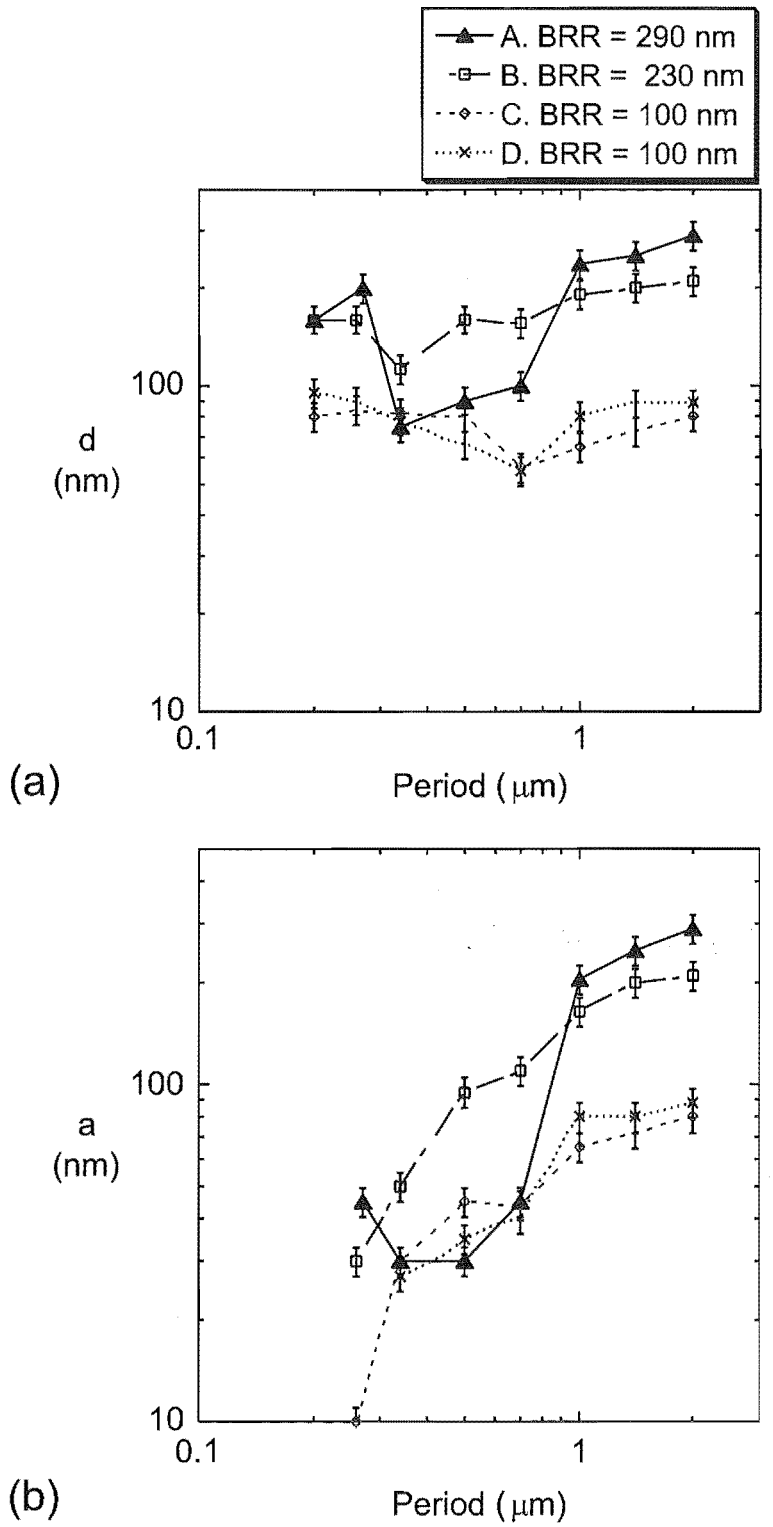


Figure 5.5 Experimental resist depth data for four samples with different exposure times, (a) plots the pattern depth d versus grating period, (b) plots the modulation depth a versus grating period. BRR refers to the background resist reduction for that exposure.

The 270 nm and 200 nm period gratings in sample A have a d similar to that of the large period gratings prior to the minimum. This indicates there has been a reduction in the light transmission through these gratings. The difference in resist profiles for large and small period gratings are highlighted in the AFM images of Fig. 5.6 and Fig. 5.7. Figure 5.6 shows a 1.4 μm period with a equal to d, as evident by the flat resist profile of the right-hand edge that is a section between two grating fingers. Figure 5.7 shows the edge of a 270 nm period grating from sample A. The modulation ($a \sim 45$ nm), from the grating pattern can be seen on top of a resist thicker than the surrounding resist ($d \sim 200$ nm), that corresponds to an unpatterned region of the mask.

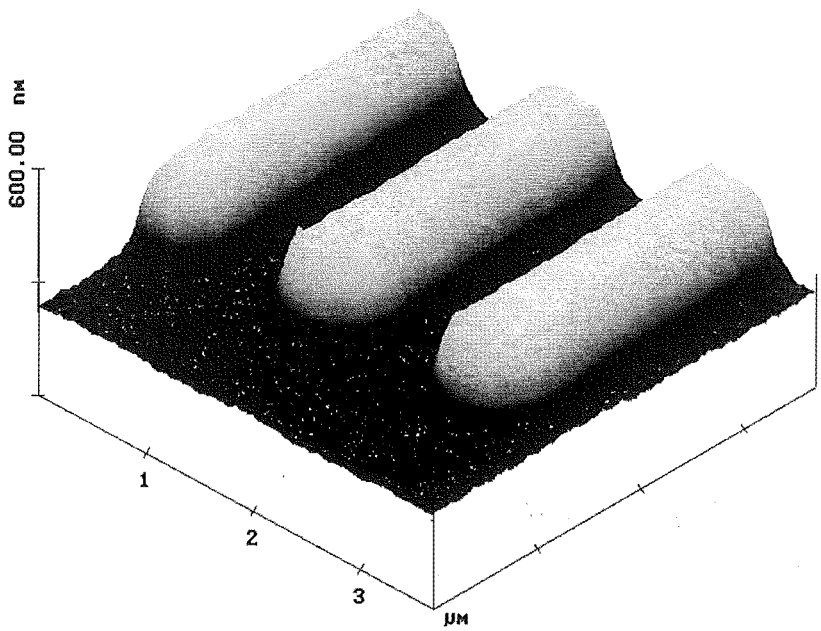


Figure 5.6 AFM image of at 1.4 μm period grating, exposed into 470 nm thick photoresist. The image illustrates that for large periods with respect to the wavelength, that $a \approx d$.

Samples C and D have similar BRR values ($\pm$ 10 nm), but differing exposure and development conditions. The results match quite closely, (within the experimental error), both exhibiting minima for d at the 700 nm-period field. Note that the minima for samples A and B occur at lower grating periods (at 270 nm- and 260 nm-period grating fields).

In Fig. 5.5(b) the resist modulation a is plotted against grating period. For large period gratings (≥ 1 μm), a and d are approximately equal, as might be expected for apertures larger than the illumination wavelength. As the grating period is reduced a trend of decreasing modulation depth is clearly evident. The grating periods from 700 nm to 340 nm for sample A show a pronounced decrease in a that deviates from the

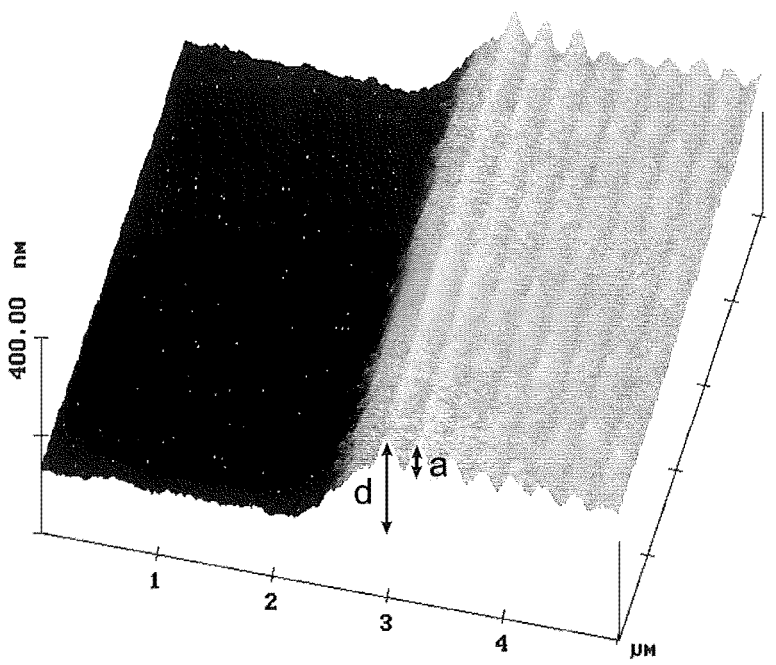


Figure 5.7 AFM image of at 270 nm period grating, exposed into 470 nm thick photoresist. The small modulation depth ($a \sim 45$ nm) compared to pattern depth ($d \sim 200$ nm) is illustrated.

prevailing trend. The reductions in a for these periods is likely to be a result of the large decrease in d owing to increased transmission through these grating fields. This equates to the top section of the resist being developed away, including part of the modulated region. Graph points for the grating periods where no discernible modulation was clear have been omitted, these are the 200 nm period for A, B, C and D, and the 260 nm period for D.

These results suggest that some care needs to be taken when using techniques similar to ENFOL for exposing grating-like structures with a variety of periods comparable to, or smaller than the exposure wavelength. Not only is there a trend of decreasing modulation with decreasing grating period, requiring thinner resist, but also differences in the transmission intensity may complicate the optimisation of exposures with different grating periods. Mask compensation techniques that add additional attenuation to appropriate fields may alleviate this problem at the cost of mask complexity. Further investigation of these effects is warranted, however the end goal of development of an additive pattern transfer process did not require this.

5.2.4 Resist Thickness and Exposure Time Versus Feature Size

During exposure trials a strong relationship between resist thickness and resolvable feature size was discernable. The thinner the resist, the smaller the patterns that could be resolved. Grating periods smaller than 500 nm could not be resolved with exposures into 120 nm thick resists. This result is consistent with the theory discussed in Section 2.2. The experiments discussed for exposures into thick resist in Section 5.2.3 also echo these findings, with a general trend of decreasing modulation depth a with grating period indicating that a decreasing resist thickness is required for full development of smaller features.

Another notable trend is that of exposure time versus feature size. Large period (*i.e.*, $> 1\,\mu$m) grating structures are resolvable at much shorter exposure times (*i.e.*, $<$ 40s). To resolve sub-wavelength grating structures, longer exposure times are required. This is echoed in Fig. 5.5(b), where a is much shallower for shorter exposure times (samples C and D) compared to the longer exposure times, (samples A and B). As the image contrast is poorer for the smaller period structures, the exposure latitude is also smaller, and an exposure time that resolves the grating adequately, while minimising the exposure to nominally unexposed regions is sought.

A comprehensive analysis of resist thickness and exposure time versus resolvable features was not undertaken due to a combination of factors. The masks were initially not robust enough to allow large-scale trials and problems with exposure variability were encountered. In the initial stages of this research, tools to enable such a comprehensive study, such as AFM, were also unavailable. However, sufficient exposures were performed to obtain reliable exposure and development conditions for the main task of developing an additive pattern transfer technique.

5.2.5 Broad-band versus Narrow-band Exposure

The majority of exposures were performed with a broad-band Hg source. Unlike projection lithography, contact lithography is not limited to monochromatic sources as it does not have the high grade lens technology that suffers from chromatic aberrations, distorting the image. There are some advantages of operating with broad-band light, most importantly interference effects due to substrate reflections and uneven resist step heights are lessened.

A more practical advantage pertinent to our system is that of exposure power. To obtain a narrow-band source requires the addition of optical band-pass filters. These significantly reduce the exposure power by reducing the number of Hg lines contributing to the exposure as well as attenuating the passed wavelength in the order of 50%.

An advantage of operating with a narrow-band source would be to reduce the exposure wavelength to one of the smaller wavelength Hg lines and so minimise diffraction

effects in this manner. A narrow-band source would also be useful for quantifying the resolutions relative to the illumination wavelength. While a thorough investigation of this was not undertaken, gratings of periods down to 270 nm were resolved with a 436 nm filter having a half power bandwidth of 9 nm and 50% attenuation at 436 nm.

In general, better pattern fidelity was obtained with broad-band exposures compared to the exposures with the narrow-band filter. This was particularly noticeable with the trilayer resist system which highlighted problems with substrate reflections.

5.3 SUBTRACTIVE PATTERN TRANSFER

A subtractive pattern transfer method using dry etching was developed in previous work, using a single layer resist system consisting of the S1805 photoresist thinned to a ratio of 1:4 [51–53]. A two stage etch is required, the first etch is a short 3 second O_2 plasma etch. This is followed by a low temperature etch using SF_6 to etch the silicon. The etch recipes are given in Table 5.2. The first etch, the de-scum, is necessary to clear a thin residual organic layer remaining in exposed regions. This undeveloped layer is a result of low exposure at the resist/silicon boundary caused by the large refractive index mismatch [83]. An anti-reflection layer would alleviate this problem but it would also increase the equivalent resist thickness as the anti-reflection layer is commonly developed along with the photoresist.

The low temperature (-100°C) etch has been found to increase the silicon etch rate [84] while etching the photoresist more slowly. Fifteen seconds is the maximum etch time before the photoresist mask is etched away and in this time the silicon is etched 120 nm. Silicon etch profiles have been shown to be improved by the addition of some O_2 [84]; this allows the formation of a sidewall passivation layer of SiO_xF_y that slows the lateral etch rate changing the profile from undercut to overcut. O_2 was not introduced to the silicon etch in this case however, due to the increase in etch rate of the photoresist that results.

Table 5.2 Etch recipes for the two stage subtractive etch of silicon.

Material	Etch Gas	Etch Rate (nm/s)	Pressure (mT)	Flow Rate (sccm)	Power Density (W/cm^2)	DC Self Bias (V)	Temp. (°C)
De-scum	O_2	-	100	50	0.62	430	20
Silicon	SF_6	8	55	100	0.62	153	-100

The smallest period grating structures that have been patterned by ENFOL, 140nm [52, 53], were transferred using this two-stage dry etch recipe. The disadvantage of this technique is that it is limited to reasonably shallow etch depths due to the short lifetime of the photoresist during dry etching. An alternative technique to alleviate

this would be the introduction of an additional barrier layer between the substrate and photoresist. A barrier layer that exhibits good etch selectivity with the substrate is desirable. There are a number of materials that satisfy this criteria, for example, SiO_2, Si_3N_4, and many metals. The more difficult problem is to find a barrier layer that shows good etch selectivity with the photoresist, and is preferably indexed matched to the photoresist to improve the exposed resist profile. These issues are addressed in the work the author has performed to develop an additive pattern transfer process in the following section.

5.4 ADDITIVE PATTERN TRANSFER

Additive pattern transfer involves the addition of a solid through a resist mask onto the substrate. Afterwards the resist is generally removed, leaving only what was deposited directly onto the substrate. The two techniques typically used for additive pattern transfer of metal films are electroplating and lift-off. This section describes the development of a lift-off metallisation process for ENFOL using a trilayer resist.

A general rule of thumb is that the maximum metal thickness that can be lifted off is approximately 50% of the resist thickness [85], although this is very dependent on the resist profile. ENFOL's requirement of ultra-thin photoresist, in the order of 60 nm, severely limits the thickness of metal that can be deposited using a single resist layer. In reality the thickness of the remaining resist can be even thinner than this for the sub-wavelength features following development as shown in Fig. 5.5. In addition, an undercut resist profile is highly unlikely from a single-layer resist ENFOL exposure. To improve the resist profile for lift-off a trilayer resist scheme is proposed. Trilayer systems consist of a bottom layer (sometimes called the planarising layer), that provides the depth for the resist profile, a barrier layer that acts as the hard etch mask for etching the bottom layer, and the top image layer. Figure 5.8 illustrates the lift-off process using this trilayer resist scheme.

The choice of material for each of the layers plays a big part in the success of the lift-off. Ideally the layers should have good adhesion and be optically well-matched to the imaging resist to optimise the exposure. Suppression of reflections is also desirable to prevent unwanted interference effects. The materials' etch properties are also important and somewhat conflicting for the barrier layer. On one hand the barrier layer needs to be etched quickly and easily with the photoresist etch mask in the first etch, yet have a high etch selectivity with respect to the bottom layer in the second etch. A number of bilayer and multilayer resist schemes have been reported [86–92], motivated by the requirement to deposit very thick metal layers, much thicker than the imaging resist, or to improve the process control of patterns over substrate topography by planarisation. For ENFOL a multi-layer resist scheme is essential for any lift-off of patterns in the order of a wavelength or below. A number of different trilayer resist systems were evaluated

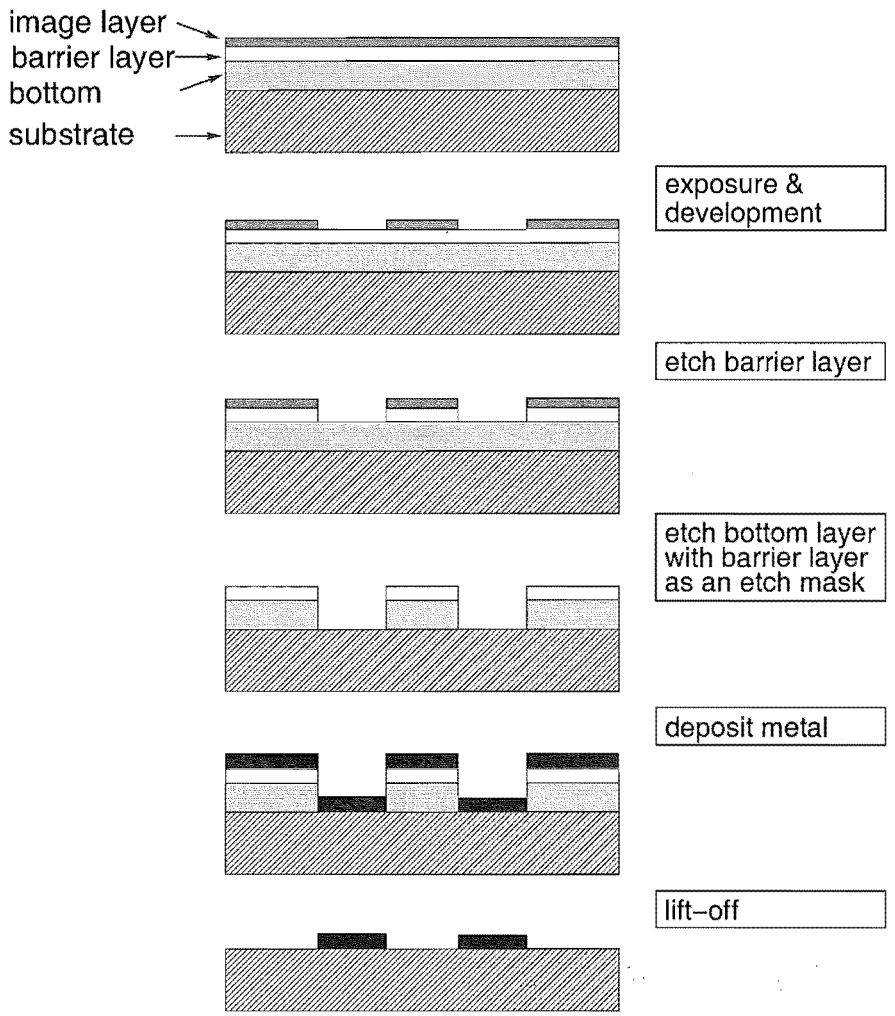


Figure 5.8 Principle of lift-off with a trilayer resist scheme.

with different materials for the bottom and barrier layers; these are discussed below.

5.4.1 Trilayer System 1: PMMA/NiCr/PR

The first trilayer system investigated was a PMMA/NiCr/photoresist combination. The PMMA was spun to a thickness of 550nm and baked at 185°C. A 15nm thick NiCr layer was deposited by physical evaporation, followed by spin application of a 120 nm thick photoresist layer (S1805 diluted 1:2 with thinner). Contact optical exposures with commercial chrome on glass masks with grating periods of 12 μm and 2 μm were initially used to develop the process, before smaller features on the membrane masks were tackled with ENFOL. A 20 s contact exposure in the Cobilt mask aligner was followed by 10 second development in MF320:DI H_2O 2:1. The next step involved transferring the resist pattern to the 15 nm NiCr film. A wet-etch process using a

standard chrome etchant $Ce(SO_4)2(NH_4)_2SO_42H_2O + HClO_4$ diluted in DI H_2O 1:1 was used to etch the NiCr. The NiCr etch rate was 90 nm/min for the thin 15 nm layer. The NiCr then acted as an etch mask during dry etching of the PMMA in an oxygen plasma at low pressure. Table 5.3 describes the dry-etch parameters for the PMMA etch.

Table 5.3 Dry etch conditions for etching PMMA with NiCr etch mask.

Material	Etch Gas	Etch Rate (nm/min)	Pressure (mT)	Flow Rate (sccm)	Power Density (W/cm^2)	DC Self Bias (V)	Temp. (°C)
PMMA	O_2	290	5	20	0.62	520	22

A PMMA profile with vertical sidewalls and ideally a slightly undercut taper is desirable for lift-off. This requires an anisotropic etch which can be obtained via reactive ion etching, described in Section 3.2. Oxygen is used to etch the PMMA as it is highly reactive with organic materials such as PMMA, converting hydrocarbons to CO, CO_2 and H_2O. A low pressure etch was used, as low as our Plasmalab 80 was capable of, to maximise the DC bias to obtain a highly vertical etch. Operating at low pressure ensures that the species have a long mean free path that reduces the probability of redeposition. Good selectivity (> 40) is obtained with the tough NiCr etch mask.

Gratings down to 700nm periods were transferred using this technique with ENFOL exposures, however this trilayer system was abandoned owing to a number of problems with the NiCr barrier layer. Firstly the lateral etch rate of the NiCr was difficult to control for small features; a faster lateral etch than vertical etch occurred, possibly owing to the formation of tough interfacial layers at the PMMA/NiCr boundary or more likely bad photoresist adhesion to the NiCr. A further disadvantage was NiCr's performance as an etch mask. While it demonstrates extremely high etch selectivity, it displays an undesirable excessively undercut profile as is illustrated in Fig. 5.9. The profile shown is of the PMMA, after the oxygen etch and prior to metal deposition for a 700 nm period grating. The photoresist has been removed during the oxygen etch and the NiCr barrier layer is visible with a large unstable overhang. While a reasonably vertical profile was produced, the undercut for small features would be too great, limiting the feature size that could be lifted-off. The undercut is thought to be caused by local heating due to eddy currents induced on the metal mask; a similar effect was reported with etching Si with a NiCr mask [84]. In addition, NiCr is poorly index-matched to the photoresist. Being a metal, electric fields tend to zero at the NiCr/photoresist boundary resulting in an underexposed region.

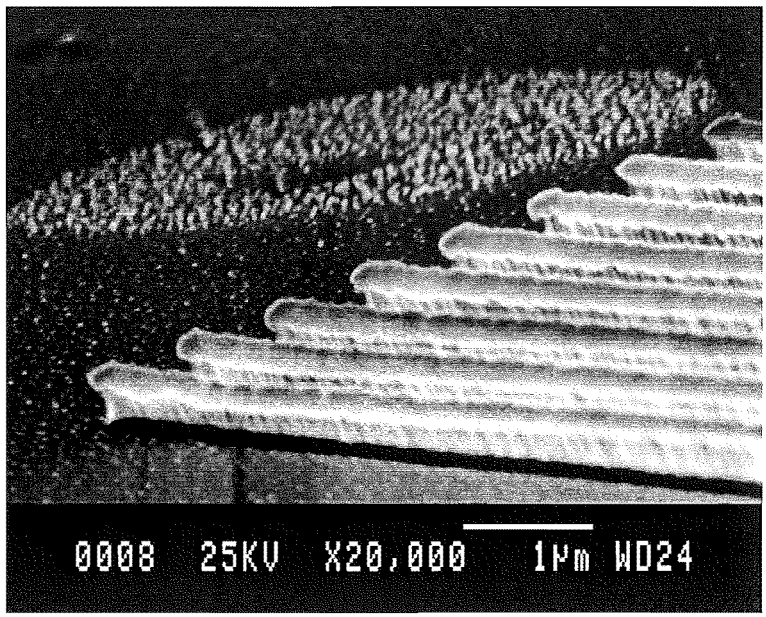


Figure 5.9 SEM image showing the profile of a the trilayer resist following the oxygen etch step for a 700 nm period grating. The sample has been Au coated.

5.4.2 Trilayer System 2: PMMA/Ge/PR

Another multi-layer scheme was trialed with the NiCr barrier layer replaced with Ge. Ge has been commonly used as a barrier layer [91] and has the advantage of being able to be dry-etched, a more controllable process than wet-etching. It is also less conductive than NiCr, reducing the undercut problem that occurs during the etch of the bottom layer. The PMMA thickness was reduced to 160 nm, as the selectivity of Ge is not as high as NiCr; about 18:1 for the PMMA etching process used. A 16 nm thick layer of Ge was evaporated onto the PMMA and then 120 nm of photoresist was spun on. Following exposure and development the samples were given a 3 s de-scum in oxygen to remove any residual resist. The Ge barrier layer was etched in CHF_3 with the recipe given in Table 5.4. Both CHF_3 and SF_6 etch Ge but CHF_3 was used as the etching of the photoresist is believed to be suppressed by adsorption of CFx monomers [93].

Table 5.4 Dry etch conditions for etching Ge with a photoresist etch mask, and for the subsequent PMMA etch with a Ge etch mask.

Material	Etch Gas	Etch Rate (nm/min)	Pressure (mT)	Flow Rate (sccm)	Power Density (W/cm^2)	DC Self Bias (V)	Temp. (°C)
Ge	CHF_3	16	2	15	0.62	523	22
PMMA	O_2	285	5	20	0.62	520	22

The bottom-layer resist profile with a Ge barrier layer did not have the same undercut as the NiCr, and appeared to have reasonably vertical sidewalls, as shown by

the 500 nm period grating in Fig. 5.10, although the thinner PMMA layer makes the undercut more difficult to determine in this micrograph. Accurate cleaving was difficult with the smaller membrane mask fields patterned with EBL and without access to a field emission SEM, assessment of resist profiles was difficult.

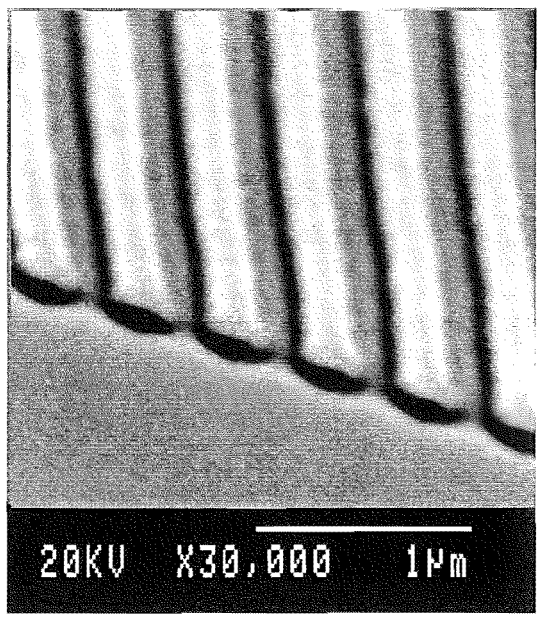


Figure 5.10 SEM image showing the profile of a the PMMA/Ge/PR trilayer following the oxygen etch step for a 500 nm period grating. The sample has been Au coated.

With this trilayer combination grating periods down to 500 nm were lifted off. Figure 5.11 shows some SEM images of gratings with 1 μm and 500 nm periods lifted off in 30 nm of Au. The Au grains are clearly visible in the micrograph. Good line fidelity is illustrated with only a small amount of line-edge roughness. The 250 nm lines and 250 nm apertures resolved in Fig. 5.11(b) is smaller than any of the illumination wavelengths contributing to the exposure, illustrating the high resolution of ENFOL. These results are also an improvement on previous reports of conformal contact lithography with Hg sources [43].

The reason why smaller patterns than those of Fig. 5.11(b) were not transferred is thought to originate from the ENFOL exposure. To improve the exposure into the resist, better index matching was sought as while the Ge barrier layer was an improvement on the NiCr, it still has a relatively high refractive index $n = 4.04 + j2.18$ at 436 nm [94], compared to the photoresist ($n \sim 1.68$). Additionally, reduction of substrate reflections is believed to improve the image contrast. This is particularly important considering the divergent nature of the illumination source.

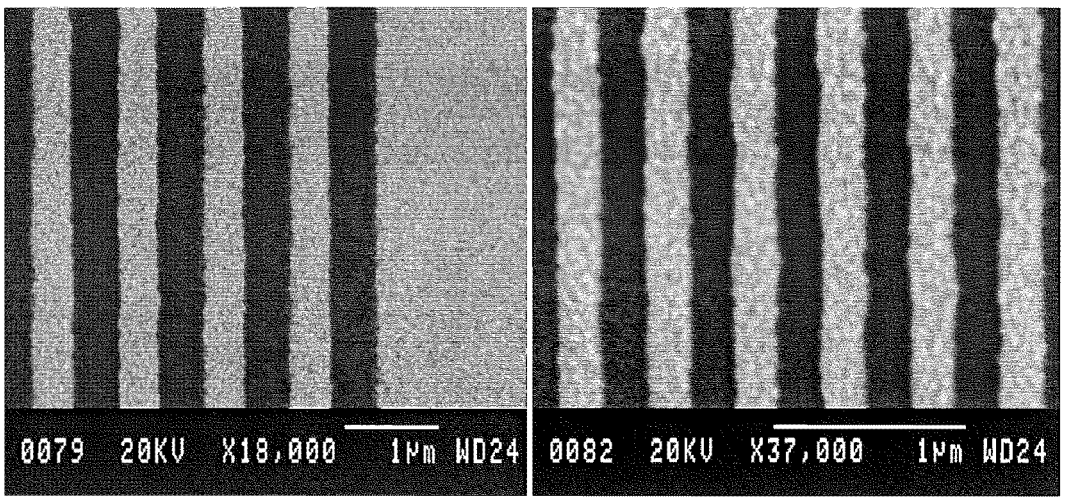


Figure 5.11 SEM images showing gratings lifted-off using a trilayer resist of PMMA/Ge/photoresist (a) 1 μm period grating and (b) 500 nm period gratings in 30 nm of gold.

5.4.3 Preventing Substrate Reflections

The large refractive index mismatch at the silicon substrate/PMMA interface causes unwanted reflections. The effect of interference can be clearly seen with a narrow-band source. Figure 5.12(a) illustrates a field with 1 μm features down to 350 nm features, following ENFOL exposure through a 436 nm filter. The pattern was transferred via etching into the underlying Ge and PMMA layers as described previously. For comparison Fig. 5.12(b) illustrates the resist pattern for the same field, exposed from the same mask but this time with broad-band illumination. The interference in Fig. 5.12(a) is particularly obvious in the pattern identifiers; the 0's appear wobbly and the numbers appear doubled. The dense patterns on the mask are very underexposed, while the sparse mask patterns overexposed. While the use of a narrow-band source had a detrimental effect on the resist images for trilayer resist systems, no obvious degradation was observed for the thin single layer resist.

The effects of substrate reflections on the resultant exposure with broad-band sources are not as noticeable as those with narrow-band/monochromatic sources. The temporal coherence of narrow-band/monochromatic sources encourages distinct interference effects that are often manifest as undesirable rings of exposed an unexposed regions. Substrate reflections are still undesirable for broad-band sources too. When exposing fields with sub-wavelength feature sizes, the image contrast is not as high as for large features and additional intensity from reflections under shadowed regions can degrade the image contrast. Obtaining a suitable trilayer scheme that minimises reflections would provide a more general purpose solution with the flexibility to pursue further work with narrow-band sources. To prevent reflections the bottom layer can

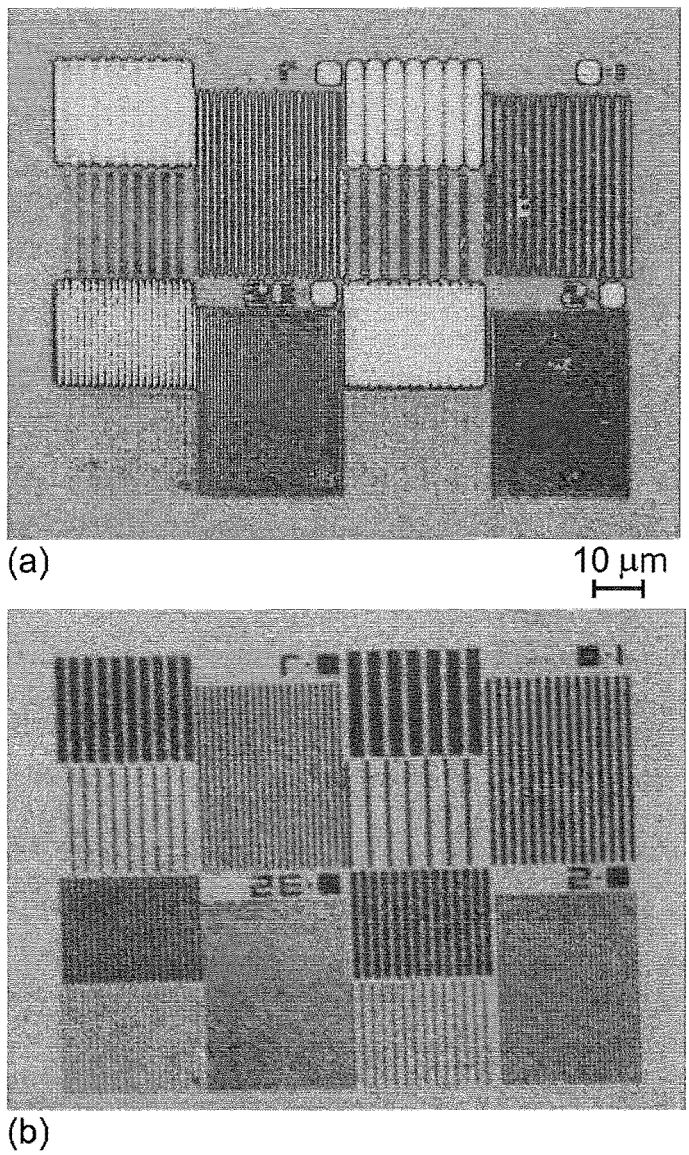


Figure 5.12 Optical micrographs comparing interference effects due to substrate reflections with a PMMA/Ge/photoresist trilayer. Two different exposure sources are used, (a) narrow-band source, the pattern has been transferred via etching into the Ge and PMMA (b) broad-band source, pattern shown in the photoresist layer.

be replaced with a highly absorbing layer. Quarter-wave matching is another solution [95], however a suitable material with a refractive index the geometric mean of the surrounding layers can be difficult to find. The thickness of the bottom-layer is also restricted to odd multiples of a quarter-wavelength. Two solutions for the absorbing layer were investigated, initially dye-doping the PMMA to increase its absorbance, and later the use of a commercial anti-reflection product.

5.4.3.1 Dye-Doped PMMA

Coumarin dyes were found to be good candidates for doping the PMMA. They dissolve readily in solvents, and a large range of dyes exist with high absorbance in the blue and UV wavelength range. Their conventional use is as the dye in dye lasers to achieve lasing at specific frequencies. Coumarin dyes have been reported previously for dye doping of PMMA [96] for reducing reflections.

The Coumarin dye 540A was trialed initially as it was available on site. The dye was dissolved in chlorobenzene solvent along with PMMA and was then spun onto a glass substrate and baked. Spectroscopy measurements, as described in Section 3.3, were performed to evaluate the transmission through the film. While this process was successful and proved the concept, the molar absorptivity was low in comparison to other dyes, so a different dye was sought to maximise the absorbance.

The Beer-Lambert law [97] approximates the absorbance A of a film of thickness u (cm) as

$$A = euc \tag{5.1}$$

where e is the molar absorptivity ($\mathrm{L\,mol^{-1}\,cm^{-1}}$), and c the concentration of the compound in solution ($\mathrm{mol\,L^{-1}}$). From tables of Coumarin dye properties [98] two other suitable candidates were found to maximise the absorbance at around 436nm. Table 5.5 gives the properties of the three Coumarin dyes tried.

Table 5.5 Coumarin dye properties. Quoted absorption maximum and molar absorptivity are for dyes dissolved in ethanol.

Coumarin Dye	Constitution	Absorption Maximum (nm)	Molar Absorptivity e (10^4 $\mathrm{L\,mol^{-1}\,cm^{-1}}$)	Molecular Weight ($\mathrm{g\,mol^{-1}}$)
540A	$C_{16}H_{14}NO_2F_3$	423	1.9	309.29
535	$C_{20}H_{19}N_3O_2$	433	5.1	333.39
521	$C_{17}H_{17}NO_3$	450	4.8	283.33

Coumarin 535 was purchased as it showed superior molar absorptivity at a more suitable wavelength, however its solubility was poor, less than 8.4 g/L compared to

40 g/L for the 540A. Another dye, Coumarin 521 was then tried, as information obtained from the Exciton manufacturer showed it had good solubility in toluene, a closely related solvent to chlorobenzene. While the absorbance peak appeared at a longer wavelength than was desirable (450 nm), the measurement was made for an ethanol solvent and using different solvents shifts the peak. From spectroscopy measurements on the 540A dye, shown in Fig. 5.13, chlorobenzene was found to give an 11 nm blue-shift. The same effect shifts the 521 absorption peak closer to the 436 nm mercury arc lamp G-line. The solubility of the 521 was greatly improved compared to the 535 dye, with 60 g/L concentrations able to be dissolved. Figure 5.14 shows the transmission through a 260 nm thick film of coumarin 521 with 5% LMW PMMA on a silica substrate. A notable dip in the transmission is observed with the minimum transmission (T) of 19% at 436 nm, confirming the required blue shift. Unfortunately the dip is relatively narrow, not taking in the I-line, and only partially attenuating the H-line. The rapid transmission fluctuations at shorter wavelengths are noise, significant owing to the low photon count as the source intensity is low here.

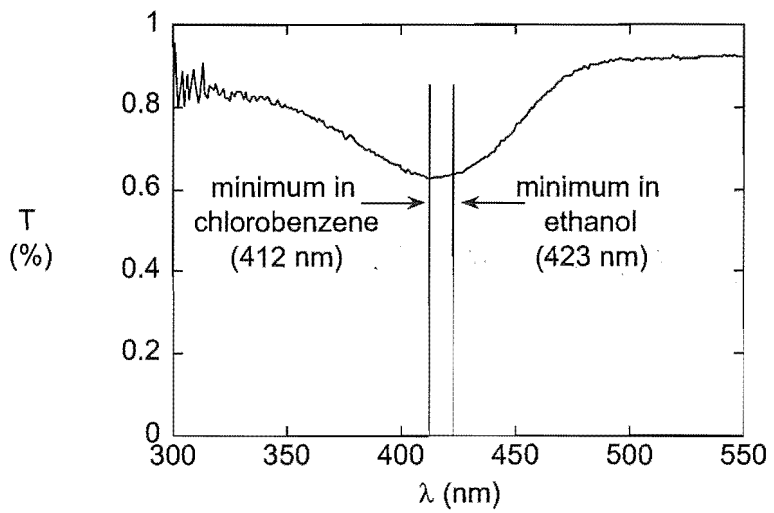


Figure 5.13 Transmission of Coumarin 540A dissolved in chlorobenzene with PMMA, showing an 11 nm blue-shift compared to manufacturers data for an ethanol solvent. The solution was initially 3% by weight dye to chlorobenzene and PMMA.

While the absorption shown in Fig. 5.14 is very good, in the bake step at 185°C the film partially crystallises, creating a surface roughness of between 400–500 nm on the nominally 260 nm thick layer. Obviously this is unacceptable for the trilayer process, so to improve the film quality the dye concentration needed to be reduced. Reducing the bake temperature also reduced the crystallisation by reducing the amount of solvent driven out of the film, although this then made the film thermally unstable causing some cracking during the thermal evaporation of the Ge barrier layer.

Successful lift-off was performed using a dye-doped PMMA/Ge/PR trilayer with

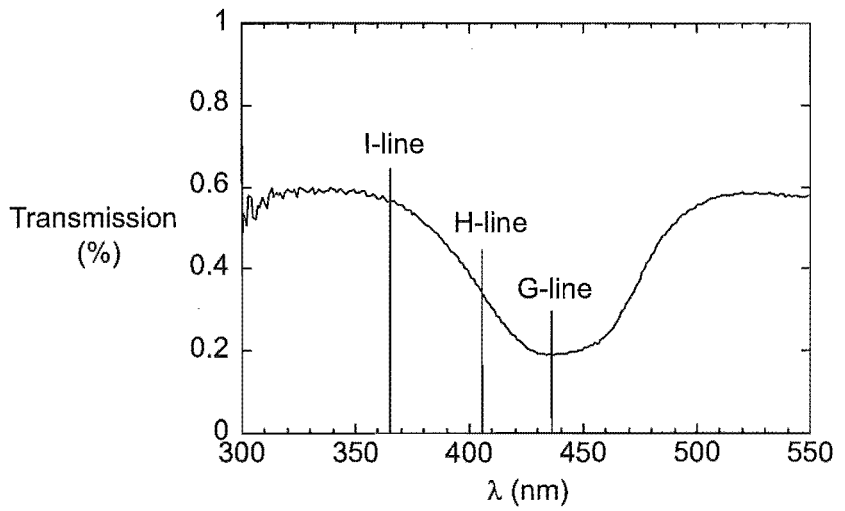


Figure 5.14 Transmission of 260 nm thick coumarin 521 with 5% LMW PMMA.

best results down to a 500nm period as illustrated in Fig. 5.15. The dye doped polymer was etched with the same recipe as in Table 5.4 but a longer etch was necessary as the etch-rate of the dye doped PMMA is lowered to 150 nm/min.

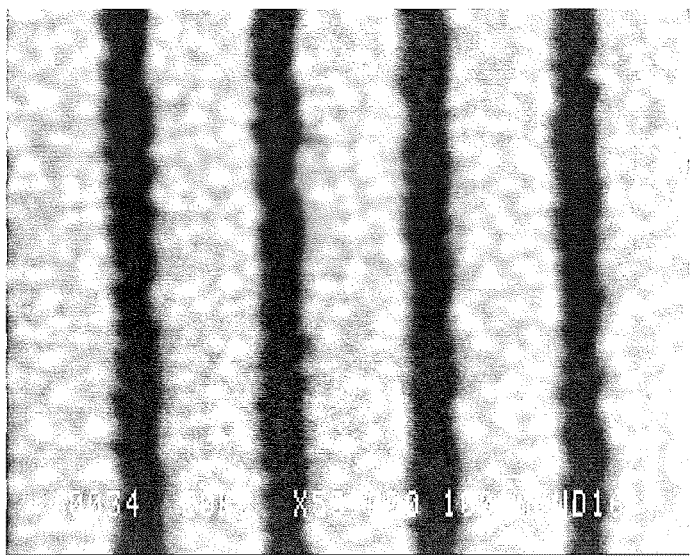


Figure 5.15 SEM image of a 500 nm period grating in 25 nm thick of Au/Pd. Trilayer resist of dye doped PMMA/Ge/PR.

5.4.3.2 Commercial Anti-Reflection Coating

A commercial anti-reflection coating (ARC) XLT, from Brewer Science was procured after the dye-doped PMMA development, and comparisons were made of the two films.

The XLT has a complex refractive index of $1.74 + i0.34$ at 436 nm and a quoted absorbance of 0.61 for 130 nm thick films [99]. It is a viscous liquid (particularly when cold) and a 2 minute spin coat at 5000 rpm was needed to obtain an even coating. The ARC was baked on a hotplate for 20 seconds at 95°C followed by a one minute bake at 170°C. The temperature of the second bake stage was reduced from that recommended by the manufacturer (186–195°C) since the film appeared mottled when baked at these higher temperatures.

The absorbance properties of the ARC were evaluated by transmission spectroscopy and compared to the dye-doped PMMA. Figure 5.16 plots the transmittance of a 120 nm thick film of the XLT, Fig. 5.16(a), compared to a 140 nm thick layer of dye-doped (C521) PMMA, Fig. 5.16(b). Both demonstrate almost equivalent absorbances for wavelengths around 436 nm with the dye-doped PMMA having a slightly lower transmission minimum. However the XLT has a broader wavelength range with high absorbance that takes in the Hg I-line and H-line as well as the G-line. A further disadvantage of the dye-doped PMMA was that, while the crystallisation had been avoided, there was still some minor cracking occurring during the thermal evaporation of the barrier layer. Based on these comparisons it was decided to continue to use the XLT ARC in further trilayer development.

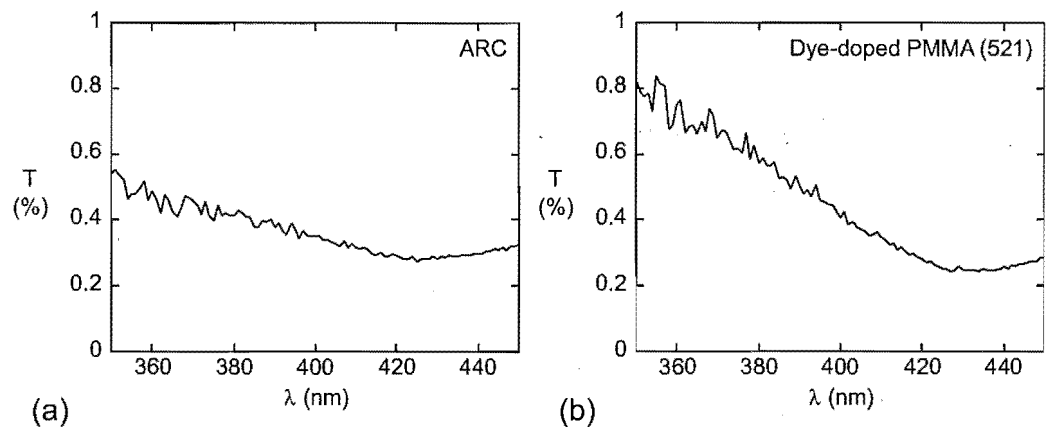


Figure 5.16 Transmission spectroscopy results for (a) 110 nm of XLT anti-reflection layer and (b) 140 nm thick dye-doped (C521) PMMA on glass substrates.

5.4.4 Trilayer System 3: ARC/SiO_x/PR

The trilayer combination that gave the best lift-off results was the Brewer XLT ARC for the bottom layer, a SiO_x barrier layer and the thinned photoresist (≈ 60 nm). SiO_x was substituted for Ge owing to its closer refractive index matching to the other resist layers. Although the exact stoichiometry of the deposited material is unknown, SiO_2 has an $n = 1.47$ at 436nm and SiO has an $n = 2.65 + 0.46$ at 436nm [100]. These values

are much closer to the refractive index of the photoresist (n between 1.6 – 1.7). The XLT was spun on as described in Section 5.4.3.2 to approximately 120nm, then a 20nm thick SiO_x layer was thermally evaporated on from SiO powder. The x denotes the uncertainty as to the stoichiometry actually obtained during the evaporation. The SiO_x was etched in CHF_3 and the XLT in O_2, with recipes similar to that for the trilayer with Ge and PMMA. The etch conditions are detailed in Table 5.6. Acetone does not readily remove the XLT ARC, so concentrated Shipley resist developer MF320 was used for the lift-off. Perhaps the only disadvantage of using the XLT over the PMMA variants was the decrease in etch rate which reduces the selectivity ratio between the SiO_x and the bottom layer.

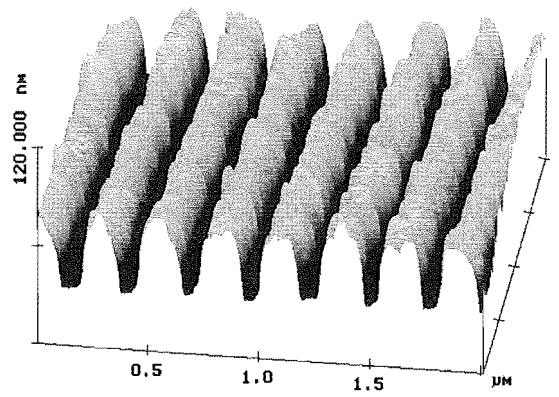
Table 5.6 Dry etch conditions for etching the SiO_x layer and XLT anti-reflection coating.

Material	Etch Gas	Etch Rate (nm/min)	Pressure (mT)	Flow Rate (sccm)	Power Density (W/cm^2)	DC Self Bias (V)	Temp. (°C)
SiO_x	CHF_3	20	3	15	0.62	506	22
XLT ARC	O_2	90	6	20	0.62	489	22

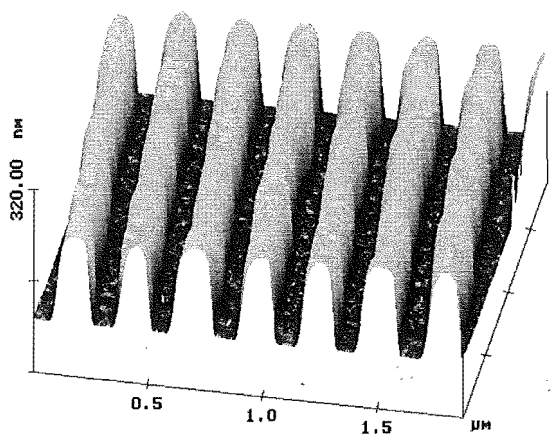
Grating periods down to 270 nm were lifted off successfully using this process. Figure 5.17 shows a portion of a 270 nm period field at each stage of the trilayer process. Figure 5.17(a) shows an AFM image of the top photoresist layer following exposure and development. Note that the photoresist is only 45 nm thick; some of the nominally unexposed resist was removed during development. While some line-edge roughness is apparent, the profile appears improved compared to resist profiles directly on silicon substrates due to improved index matching between the resist and the underlying layer (SiO_x in this case).

Following the two stage dry etch of the SiO_x and ARC, (recipe given in Table 5.6), the profile has been greatly improved as seen in Fig. 5.17(b). The height of the features is now 150nm, improving the aspect ratio of the nominally 135nm wide lines from 0.33:1 to better than 1:1. A 30 nm NiCr film was then deposited and lifted-off in MF320, as shown in Fig. 5.17(c). The line-edge roughness present in the initial resist image appears to have been transferred through to the metallised pattern. The metal lines appear broadened suggesting there may be some undercutting during the dry etch stage.

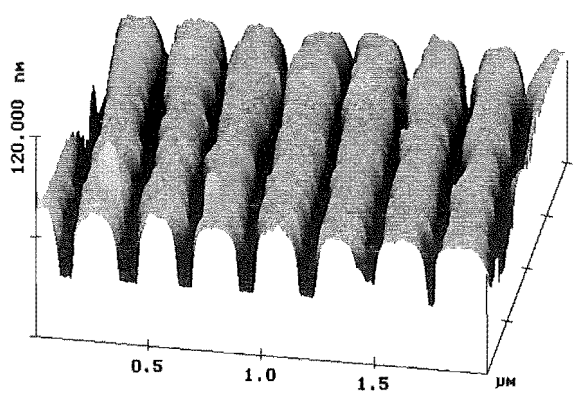
The reason why grating periods smaller than 270 nm were not transferred with the additive pattern transfer process was not a function of the trilayer process, rather these gratings were not resolved in the photoresist. It is not well-understood why patterns were able to be resolved down to 140 nm in the single-layer resist and subsequently etched into silicon [52,53], but not with the trilayer resist. The improved index matching of the photoresist with the trilayer process compared to the single layer resist on silicon



(a) Photoresist following exposure and development.



(b) ARC and SiO_x following dry etch.



(c) Lifted off metal

Figure 5.17 AFM images of a 270 nm-period grating, (a) in photoresist following exposure and development (note it is only 45 nm thick), (b) in ARC and SiO_x illustrating the improved profile following dry etching, (c) in 30 nm of NiCr following lift-off.

would suggest that the trilayer resist would be more likely to have improved exposures into the resist.

While the exposure and pattern transfer with the single layer resist was repeatable for small periods, the results were limited to only one mask. This suggests that the particular mask profile may have had a strong influence. A likely possibility is that the exposures for the 200 nm and 140 nm period gratings were made through fields on the mask that did not lift off. These mask fields would appear as two gratings, one on top of the other, but offset from each other. Partially polarised light would pass through the first grating (on top of the membrane) before being transmitted through the second grating, separated by the thickness of the PMMA. Image contrast due to mechanically imprinting the resist can also not be discounted, however no evidence of an image reversal that would result from imprinting has been observed. Further research is anticipated in this area.

5.4.5 Line-Edge Roughness

Line-edge roughness (LER) is evident in the photoresist following development of an ENFOL exposure, as shown in Fig. 5.18. It is not only an issue for ENFOL, but is currently of concern for the majority of lithography technologies that rely on exposure of a chemical resist, such as EUV, projection lithography and electron beam lithography [3, 101]. LER is an increasing problem as the linewidth is reduced and the roughness becomes proportionately more significant. For many applications this roughness is unacceptable, for example, for current carrying wires electromigration effects can be worsened and for transistor gates, switching performance could be compromised.

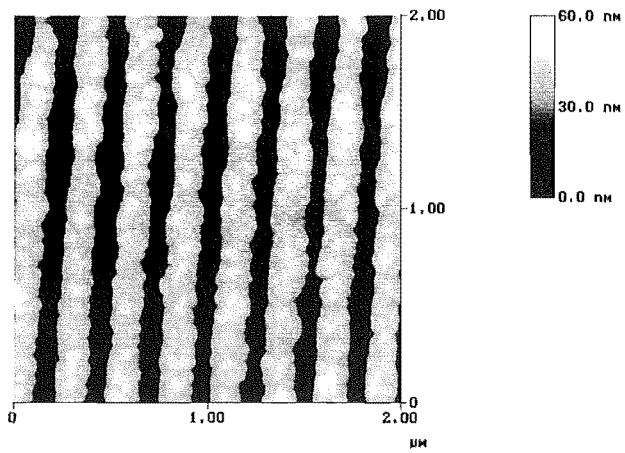


Figure 5.18 AFM image of resist lines on a 270 nm period. The edges of the lines show some roughness and the granularity of the resist is visible.

There are a number of factors that contribute to LER in the ENFOL process.

Firstly the resist is inherently grainy. AFM measurements suggests that the grain size of the resist used here is in the order of 40nm, although these measurements are only an indication, owing to the uncertainty of the tip radius. The discrete nature of the grains prevent a smooth line being formed. The S1805 resist was not designed for these high resolutions and LER issues have been acknowledged by the manufacturer [102]. The problem is amplified in the ENFOL work due to partial exposure through the absorber. The NiCr absorber used on the masks is only 35–40 nm thick, which is only a tenth of the exposing wavelength. LER may be improved by a thicker mask patterning layer, although some partial exposure is unavoidable at the resist edges.

The integrity of the thinned $\approx$ 60 nm resist was investigated using the AFM. An exposure was taken with a traditional chrome on glass mask and developed, then images of the resist from a centrally unexposed region were compared to resist at the edge of an unexposed region. For approximately 1 μm from the resist edge exists a well defined thickness gradient between the unexposed areas on the mask and exposed areas. In this region, quite high surface roughness was visible, as shown in Fig. 5.19, which could contribute to LER. Note some element of AFM artifact is suggested by the uniform alignment of the resist grains. The resist in the unexposed area was smooth and homogeneous. These findings echo what has been observed when comparing resist shadowed by the silicon supporting region having had no exposure, to resist features shadowed by the NiCr and partially exposed in the ENFOL exposure.

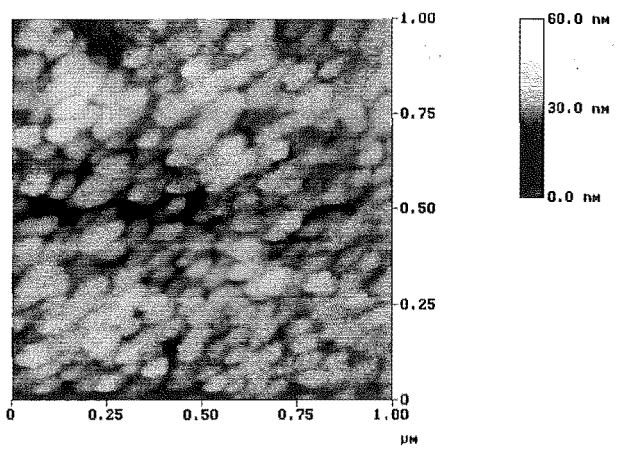


Figure 5.19 AFM image of partially exposed resist on the boundary between the exposed and unexposed region with a chrome on glass mask.

Another possible source of LER may be the non-flat profile of the metal patterns on the mask, as discussed in Section 4.5. During the exposure, the torn edges of the metal absorber may get flattened under vacuum pressure to create a rough absorber line-edge. Sharp metal discontinuities can also create localised areas of high electric field resulting in non-uniform exposures.

To improve LER, replacement of the thinned but otherwise conventional novolac based resist will be required. Ideally an alternative resist that combines high sensitivity with good resolution and a small grain size is desirable. Due to the requirements of a thin resist layer, it must also exhibit reasonable robustness with dry etch processes, to enable pattern transfer. Surface imaging resists and self-assembled monolayers are alternative resists that provide a thin imaging layer, however no solution as yet incorporates all of the desirable properties listed above. Goodberlet *et al.* has demonstrated small LER using a PMMA resist at shorter wavelengths [46,47].

5.5 RESOLUTION

The smallest grating period that was transferred with the trilayer lift-off metallisation process was 270 nm. Table 5.7 provides some comparison of the grating resolutions obtainable with optical projection lithography at wavelengths equivalent to the dominant spectral lines of a mercury arc lamp. In column two, the minimum resolvable period for an optical projection system, using Equ. 2.5 from Section 2.2 is calculated, for normally incident light, and assuming the best case condition of NA = 1. Column three contains the grating periods equating to cut-off (see Section 2.6), below which all diffracted light will be evanescent. These are calculated for a grating in photoresist with a refractive index of 1.6.

Table 5.7 Minimum resolvable grating period with optical projection lithography, and cut-off grating period below which entails a fully evanescent exposure versus relevant wavelengths of a Hg lamp. Normally incident light is assumed in both cases.

Wavelength (nm)	Optical Projection Min. Period (nm)	Cut-off Grating Period (nm)
365	365	228
405	405	253
436	436	272.5

It is clear from the table that the minimum transferred period of 270 nm is well below that of an equivalent optical projection scheme, 1.6 times smaller at a wavelength of 436nm and 1.35 times smaller than at a wavelength of 365nm. This demonstrates sub-diffracted limited resolution with respect to conventional projection techniques. With ENFOL exposures and subtractive pattern transfer, gratings down to 140 nm periods have been reported [52,53], suggesting that solely evanescent fields were responsible for the exposure contrast. This has not been demonstrated with additive pattern transfer although the 270 nm periods are just within the range of the cut-off grating periods for the relevant Hg lamp lines.

5.6 SUMMARY

ENFOL experiments were conducted with a number of resist systems. Exposures into thick photoresist illustrate the strong relationship between the required resist thickness and grating period, by showing a strong decrease in the modulation depth as the period is reduced. Differing transmission through fields with different grating periods were highlighted for periods close to, and below the exposure wavelength. A notable increase in the transmission was observed for periods in the order of the illumination wavelength. This could have a negative impact when exposing the equivalent of fields of different periods, as the optimal exposure–development conditions may differ. The addition of attenuating layers to the mask would circumvent this problem at the cost of mask complexity. Large period fields in the thin resist have greater process latitude with respect to exposure time.

A lift-off metallisation scheme compatible with ENFOL was developed utilising a trilayer resist. The combination of a commercial anti-reflection coating (XLT) as the bottom layer, SiO_x barrier layer and photoresist imaging layer was found to give the best results. The anti-reflection coating reduced the effect of interference from substrate reflections. Improved index matching between the different resist layers optimised the photoresist profile resulting from the initial exposure. Anisotropic etching of the SiO_x and anti-reflection coating created steep sidewall profiles compatible with lift-off requirements.

An anti-reflection bottom layer was also developed by dye-doping PMMA with a laser dye, Coumarin 521. This film exhibited low transmittance ($\approx$ 25%) around 436 nm. While this anti-reflection coating was dropped in favour of the commercial XLT product, prior to this, it was used successfully in the trilayer resist for lift-off metallisation.

The smallest period gratings that were lifted-off were 270 nm, which equates to a resolution better than 1.6 times smaller than that achievable with a G-line projection system with normal illumination. This grating period is also right on the edge of the cut-off regime, where, for periods below this, exposure contrast is attributable to solely evanescent components of the exposure field.

ENFOL exposures exhibit some degree of line-edge roughness. Mask patterns with flatter profiles and thicker metal layers as well as improved collimation may improve line-edge roughness. For improved results at the nanoscale level, a higher resolution resist having good contrast and small grain size will be essential.

Chapter 6

NEAR FIELD SIMULATIONS

6.1 INTRODUCTION

A method of modelling the evanescent near field is necessary to aid understanding of the exposure mechanism and to identify what may limit the achievable resolution of ENFOL. Modelling has a number of advantages over experimentation. Parameters often difficult to vary experimentally are easily within the control of the modeller. Additionally simulations can typically achieve a faster turn-around of results, with less ambiguity. On the other hand, it is important to keep in mind the difficulty of obtaining a model identical to the experimental reality. The intended outcome of such modelling is to guide the experimenter to select a smaller number of well chosen experiments to undertake.

This chapter begins by describing the various numerical techniques available for solving diffraction problems rigorously. A description of the Multiple Multipole Program that is used to simulate the ENFOL system is provided and the process of setting up a model described. A simplified ENFOL model consisting of a Cr grating suspended in air is initially constructed. This enables evaluation of the characteristics of the evanescent near-field without the distractions of secondary effects due to interactions with multiple layers. From this simple model, a range of parameters were varied. The effect of decreasing grating period is evaluated which strongly impacts on the depth of field of an image transmitted through the grating. The effect of parameters such as grating duty cycle, the radius of curvature of absorber corners, grating thickness and grating conductivity is also observed. A more realistic ENFOL model incorporating multiple layers including the mask material, photoresist and silicon substrate is also discussed.

6.2 RIGOROUS COMPUTATIONAL METHODS

While electromagnetic diffraction problems are mathematically well-defined in terms of Maxwell's equations and boundary conditions, they are also very difficult to find solutions for. Only very simple geometries such as Sommerfeld's diffraction of a plane wave

by an infinitely conducting half-plane have been solved rigorously analytically [103]. To solve more complicated geometries, simplifying approximations have been necessary. Early physical-optics theories, such as Kirchoff's diffraction theory and the Huygen-Fresnel theory have yielded closed form solutions to more complicated diffraction problems via simplifying approximations. These theories have provided valuable insight, but it is important that the approximations are understood, in order that the range of validity of the solutions is known. Both Kirchoff's and Fresnel's theories are scalar methods and as the sizes of the features that diffract light shrink to the same order as the wavelength, the approximations used cease to be appropriate. Variations in the boundary conditions due to the different polarisations become significant and results become inaccurate. Infinitely conducting screens are assumed, that are infinitely thin and yet look infinitely absorbing so that there is no reflection. To ascertain the field at any particular point the field at the diffracting aperture needs to be known and an additional simplifying assumption is that the field at the diffracting aperture is equal to the incident field and the fields at the conducting screen are zero. These assumptions are reasonable a long way from the diffracting aperture but in the near-field region they are invalid.

Lin was one of the early pioneers to investigate the near field for photolithography applications. He solved the diffraction problem of a slit formed by a gap in two infinitely conducting, infinitely long half-planes [104]. Lin modified Sommerfeld's exact solution by superposing the solution for each half-plane and incorporating a simplified two-order approximation to take into account the interaction between the two half-planes. While Lin's method is not strictly rigorous, useful results were obtained by this method that is valid for slit widths down to half a wavelength. Results have been presented for one, three and ten wavelength slit widths, initially at normal incidence and then at varying incidence. His results for the fields resulting from exposure with TM polarised light have a similar form to the results presented for gratings later in this chapter. The TE results show strong peaks at the exit aperture of the slit with the number of peaks equal to the number of wavelengths of the slit width. Lin demonstrates that these results converge to Fresnel and Fraunhofer approximate theories for large slits. Despite the trivial appearance of the geometries, analytically these are difficult problems to solve.

This method is not suitable for analysis of ENFOL for a number of reasons. Firstly, the analysis is limited to slit geometries. Secondly, it does not incorporate realistic material properties for the conductors, and nor does it consider conductors of finite width. Thirdly, Lin's analysis is limited to slits of half a wavelength, whereas for ENFOL features with much smaller dimensions than this are of interest.

The vast gains in computing power over the past three decades have enabled the development of powerful numerical techniques. With these, increasingly more complex diffraction problems have been able to be solved rigorously, providing new information to understand the near field. A number of the more common rigorous numerical meth-

ods are summarised below with some discussion of their advantages and disadvantages with respect to the ENFOL problem. They can be loosely categorised as differential methods or integral methods, where the differential methods are based around solving the differential form of Maxwell's equations and similarly the integral methods solve the integral form of Maxwell's equations.

6.2.1 Finite Difference Time Domain (FDTD) Methods

The finite difference time domain (FDTD) method is becoming increasingly popular for solving electromagnetic problems. One of its main attractions is its simplicity. FDTD solves Maxwell's time dependent curl equations directly,

$$\nabla \times \boldsymbol{E} = -\mu \frac{\partial \boldsymbol{H}}{\partial t}, \tag{6.1}$$

$$\nabla \times \boldsymbol{H} = \sigma \boldsymbol{E} + \varepsilon \frac{\partial \boldsymbol{E}}{\partial t}. \tag{6.2}$$

The region to be analysed is discretised into two interleaving grids. One grid contains the nodes where the electric field is solved, and the second grid is made up of the nodes where the magnetic field is solved, as illustrated in Fig. 6.1. A first-order central difference technique is employed,

$$\frac{1}{A}[E_{z1}(t) + E_{y2}(t) - E_{z3}(t) - E_{y4}(t)] = -\frac{\mu}{2\delta t}[H_{x0}(t+\delta t) - H_{x0}(t-\delta t)] \tag{6.3}$$

where A is the area of the face made up by the nodes given on the left of Equ. 6.3 and δt the time step. The electric field components are known from the previous calculation at time t, as is $H_{x0}(t-\delta t)$, this leaves $H_{x0}(t+\delta t)$ as the only unknown to be solved for. In the next iteration the electric fields are solved for in the subsequent time step. This progression of the calculation is known as the leap-frog technique, where the electric and magnetic fields are solved at alternating time steps.

Clearly an initial field must be provided at the start of the calculation. The calculation then iterates until it converges to a steady state solution for the EM fields. Being a time domain technique it is particularly well-suited for transient analysis problems. A zero field is initially specified, and then an excitation introduced. Incorporating materials that have time or exposure dependent properties should also be possible. This could be useful for looking at photoresists that undergo "bleaching" during an exposure, *i.e.*, the photoresist absorption decreases with exposure. While this type of behaviour could be incorporated into time domain codes such as FDTD, the time-varying behaviour is slow compared to the time variation of the fields and may not be the most efficient method of evaluation.

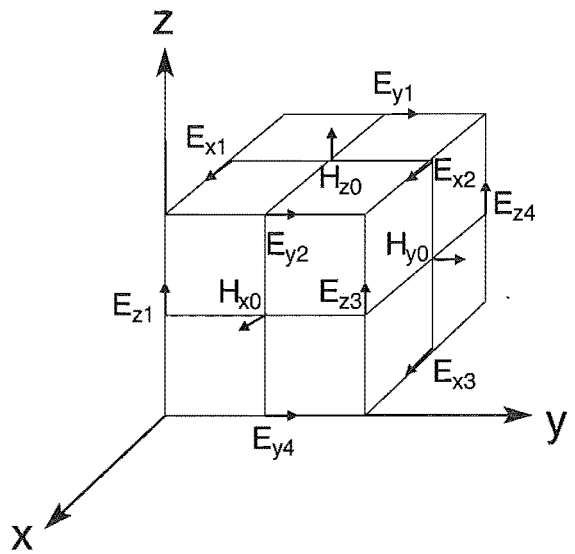


Figure 6.1 Basic unit element of the FDTD lattice, illustrating interleaved electric and magnetic grids.

The size of the grid has a direct influence on the accuracy of results and must be chosen to be significantly smaller than the finest structure on the problem geometry. The computation time and storage is directly proportional to the volume of the solution space, which means that there is the common trade-off between accuracy and computation time. A disadvantage of this technique is that problems can get unwieldy with the combination of large areas and fine structure. However recently codes have been developed that take into account axial symmetries and allow for non-uniform grids [105] to improve computational efficiency. FDTD has been applied to near-field problems such as scanning near-field optical microscopy [106] as well as for optical lithography applications, for sub-wavelength slit geometries [107], and phase mask structures illuminated at oblique incidence [108].

6.2.2 Finite Element Method

The finite element method (FEM) is an integral method; like FDTD it is a domain method where the volume to be analysed is broken up into a mesh of elements. These elements can be chosen to be of different sizes, smaller around geometric detail and larger elsewhere. It is a frequency domain method and generally works by solving for the unknown fields by minimising an energy functional. The energy functional is defined by Hubing as "an expression describing all the energy associated with the configuration being analysed" [109]. It includes the energy stored in electric and magnetic fields as well as conduction losses. The requirement of a meshing algorithm results in additional complexity, in particular very small meshes are required in highly conductive regions

to take into account the rapid changes of the fields in the metal. Calculation times are typically long. A further disadvantage of this method is that only closed geometries can be simulated. Some recent implementations have worked around this problem by introducing special radiation boundaries to allow open geometries to be simulated [110].

6.2.3 Rigorous Coupled Wave Analysis

The rigorous coupled wave analysis (RCWA) formulated by Moharam and Gaylord [111] solves the diffraction problem of periodic one-dimensional (1-D) and two-dimensional (2-D) gratings. It is a differential method that expresses the field inside a grating region in terms of modal expansions that are solutions to Maxwell's equations. These are then rearranged to a state-space description and solutions found for the state variables. The grating is constructed layer by layer so that the permittivity does not vary in the y-direction, as shown in Fig. 6.2. Each layer looks like an individual lamellar grating. A solution for the state variables is found in each layer with the coefficients determined by consideration of the boundary conditions. The field outside the grating can then be determined via phase matching.

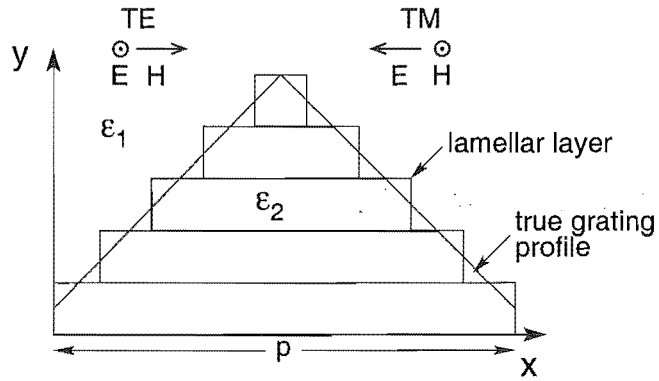


Figure 6.2 Schematic representation of grating profile using the RCWA method. The linear polarisations TE and TM are also defined.

While RCWA has been shown to be very efficient for solving grating problems [111] it converges slowly for 1-D metallic gratings illuminated by TM polarisation [112], the subject of interest of the ENFOL problem. Solving highly conductive gratings illuminated by TE polarised light is a mathematically simpler problem than TM polarisation. TE polarisation results in scalar wave-equation for the electric fields inside the grating as opposed to the vector wave-equation that results for the TM polarisation. There is also a discontinuity in the E_x field for the TM polarisation that requires a high number of diffracted orders to be maintained in the calculation to obtain convergence, consequently requiring large amounts of memory and long computation times. A reformulation of the RCWA method has been reported to improve this slow convergence [113].

RCWA is limited to the solution of infinitely periodic structures. Grating profiles are restricted to being stepped, staircase structures, as each grating must be made up of the lamellar strips. When simulating smooth profiles, this can result in significant errors. Using large numbers of very thin lamellar can improve this, however it is computationally intensive. A commercial package based on the RCWA method is currently available called GSolver [114], and this has been used for the validation of near-field simulation results presented in this thesis.

6.2.4 Green's Tensor Technique

The Green's tensor technique is an integral technique that has been used for investigating electromagnetic scattering in polarisable backgrounds [115, 116]. The problem geometry is separated into two, a background geometry that can either be a homogeneous material or perhaps a multi-layered background, and localised scatterers that may be protrusions or inclusions. The problem is solved by application of

$$\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}^0(\boldsymbol{r}) + \int d\boldsymbol{r}' \boldsymbol{G}^B(\boldsymbol{r}, \boldsymbol{r}') k_0^2 [\varepsilon(\boldsymbol{r}) - \varepsilon^B(\boldsymbol{r})] \boldsymbol{E}(\boldsymbol{r}') \tag{6.4}$$

where $\boldsymbol{E}(\boldsymbol{r})$ is the field scattered by the system with a permittivity $\varepsilon(\boldsymbol{r})$ when illuminated by an incident field $\boldsymbol{E}^0(\boldsymbol{r})$ propagating in a background with permittivity ε^B. $\boldsymbol{G}^B(\boldsymbol{r}, \boldsymbol{r}')$ is the Green's tensor associated with the background, *i.e.*, the field resulting from a point-like source excitation and k_0 is the vacuum wavenumber.

An advantage of this system is that only the localised scatterers need to be discretised to solve the problem, which is more efficient than other meshed techniques such as FDTD and FEM. It is able to investigate scattering from objects of sub-wavelength size and is suitable for finite structures where the outer boundary conditions are accounted for in $\boldsymbol{G}^B(\boldsymbol{r}, \boldsymbol{r}')$. This technique was used for simulations to investigate exposures through finite light-coupling masks [49, 50] a conformal contact lithography technique described in Section 1.3.4.3.

6.2.5 Generalised Multipole Technique (GMT)

The generalised multipole technique (GMT) is another integral technique. It is a boundary method so only the boundaries need to be discretised, and it operates in the frequency domain. In this approach, known analytic solutions (called expansions) of the field equations are used to expand the fields in the different domains of the geometric problem. For statics problems this equates to solving Laplace's equation

$$\nabla^2 \boldsymbol{U} = 0, \tag{6.5}$$

and for time-harmonic problems this equates to solving the Helmholtz equation

$$\nabla^2 \boldsymbol{U} + k^2 \boldsymbol{U} = 0, \tag{6.6}$$

where $\boldsymbol{U}$ is the field of interest, either electric or magnetic, and k the wavenumber.

The expansions are based on sine, cosine, Bessel, Hankel and Neumann functions. These act as fictitious sources for the electromagnetic (EM) fields. A commonly used expansions is the multipole, hence the technique's name. There are a number of techniques with a similar flavour to the GMT, and often it is in the choice of expansions that they differ. Some techniques for example, are restricted to using only monopole expansions [117].

The boundaries are discretised by the placement of matching points. An overdetermined set of equations can then be formed comprising the boundary conditions at these matching points. The expansion coefficients are then numerically determined by calculating the fields at the matching points to obtain the least mean square error in the boundary conditions. A generalised point-matching technique is also applied which introduces weighting functions to reduce the dependency on the choice of matching points. Once the coefficients are determined, the field can be calculated at any point.

An advantage of GMT is that as only the boundaries are discretised, this can often result in a smaller problem to be solved in comparison to volume discretised techniques such as FDTD and FEM. The disadvantage of GMT is that there is no reliable algorithm for the placement of the expansions. Heuristic knowledge is required to place the expansions, which takes time and experience to acquire, particularly if high accuracy with reasonable computation time is to be attained. The computation time is directly related to the number of expansions, the order of the expansions and how finely the boundary has been discretised.

The Multiple Multipole Program (MMP) is an implementation of the GMT that implements the generalised point matching method and multiple multipoles as sources. This was the primary resource used in simulations by the author for evaluating the near-field for various ENFOL-like geometries.

6.3 MULTIPLE MULTIPOLE PROGRAM

The Multiple Multipole Program (MMP) is a GMT based technique developed by Christian Hafner from the Swiss Federal Institute of Technology in Zurich [118–120]. The particular implementation used, MaX-1 [121], was also developed by Christian Hafner. MaX-1 is a two-dimensional (2-D) solver which incorporates both an MMP and a Generalised Finite Difference solver, but only the MMP solver was used. As the program is two-dimensional (2-D), in the third dimension (z), translational symmetry

is assumed. For the gratings simulated with their cross-section in the x-y plane, the conductors extend in the z direction, appearing infinitely long.

The first step in the simulation is the definition of the model. The model consists of three main components - domains, boundaries and expansions. A simple two-domain geometry is depicted in Fig. 6.3.

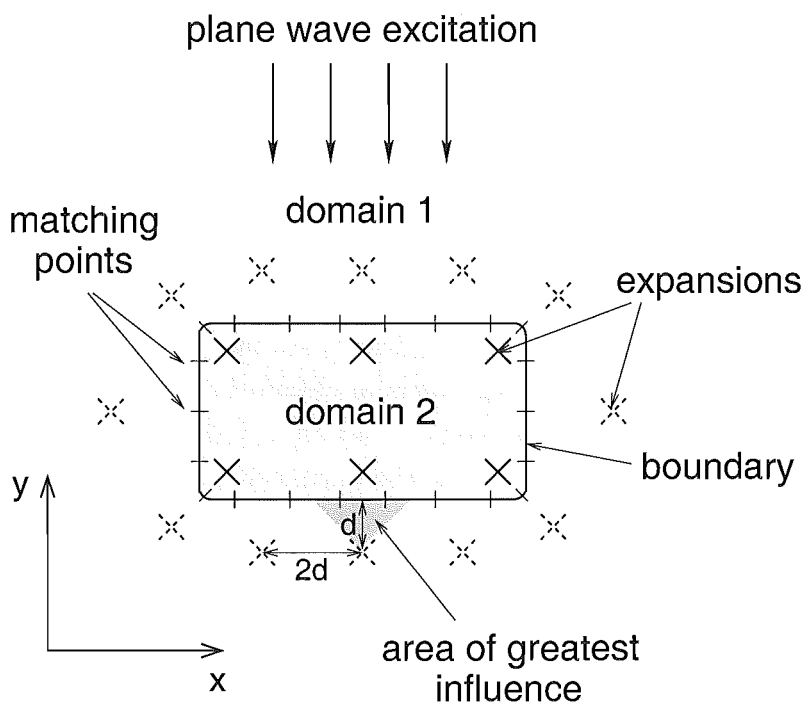


Figure 6.3 Schematic showing the constituent components of the MMP model.

6.3.1 Domains

Domains are regions with homogeneous, isotropic material properties, with their area defined by boundaries, to separate the different domains. The electromagnetic material properties are specified in terms of a complex permeability and complex permittivity, relative to free space. Values for the permittivity and permeability may be fixed or formulae may be specified to incorporate dependent variables that may be varied during a batch run of simulations. The ability to apply a formula for the material properties has been useful for incorporating the frequency dependent nature of materials when simulating the same geometry with source excitations of different frequencies.

Material parameters used in the results presented here were obtained from tables of optical properties [94], [122]. These tables generally quote material properties in terms of their complex refractive index n, where $n = n' + in''$, n' is the real refractive index and n'' the extinction coefficient. From these a complex permittivity is determined

from Equ. 6.7.

$$\varepsilon = n'^2 - n''^2 + i2n'n'' \tag{6.7}$$

A library of domains with different material properties can be created in MaX-1. The provision for the simulation of a perfect electric conductor with infinite conductivity is allowed, having the reserved domain number 0. The perfect electric conductor was not used in these simulations for two reasons. Firstly, at optical frequencies metals have an appreciable skin depth, making the investigation of finite conductivity metals more realistic. Secondly, the singularity that occurs with gratings illuminated with TM polarisation resulted in large errors.

6.3.2 Boundaries

Boundaries are the lines which separate the different domains as shown in Fig. 6.3. They are created using c-polygons [118], which is a flexible method by which almost any geometry can be created. Boundaries are linked to their associated domains by specifying the domain number on both sides. A number of types of boundary are provided by MaX-1 with each boundary type implying the boundary conditions to be applied in the calculation, amongst other things. The boundary types are described in Table 6.1.

Table 6.1 Description of the types of boundary available in MaX-1.

Boundary	Definition
Usual	Default that lets MaX-1 select the boundary conditions to apply.
SIBC	Applies a surface impedance boundary condition, for good but not perfect electric conductors.
$X-periodic$	Indicates a periodicity in the x-direction.
$Y-periodic$	Indicates a periodicity in the y-direction.
Specific	Allows the user to specify the particular boundary conditions to be matched, *e.g.* the transverse E or H field, the z component of E or H, or the normal components of D or B.

As MMP is a boundary method, it is the boundary that is discretised as opposed to the domains. Figure 6.3 illustrates how the boundary between domain one and two might be discretised using matching points where the boundary conditions are evaluated. The discretisation is an automated process, however the user guides it to some extent by specifying variables in the "matching point definition" area of the MMP dialog box in MaX-1. It is important that enough matching points are defined to ensure the spacing between the points is small compared to the effective wavelength in the medium, and small compared to the problem geometry. Enough points need to be defined to meet the criteria of an over-determined system of equations.

6.3.3 Expansions

The core of MMP is the definition of the set of expansions for a model. The expansions act as fictitious sources of the EM fields. An alternative GMT technique is known as the method of fictitious sources [123]. There are a variety of basis expansions, namely multipoles, Bessels, Rayleigh waves, plane waves and harmonic expansions. A brief description of these expansions follows, and further details can be found elsewhere [118].

Multipoles consist of Hankel functions, a type of radial function that is a linear combination of Bessel and Neumann (Bessels of the second kind) functions. These have a singularity at the origin and therefore must have their origin in a different domain to the one in which they are active. This is illustrated in Fig. 6.3 where the expansions with dashed crosses define the fields in domain two (grey) and similarly the expansions with solid crosses define the fields for domain one.

Bessel expansions are also a type of radial function but they don't have a singularity at the origin. They are located in the domain of interest and are most appropriate for modelling the EM fields for circular shaped geometries. They are restricted to closed geometries as the functions tend to infinity at large radial distances.

Rayleigh waves are expansions dependent on the periodicity of the grating. They represent the fields diffracted from a grating. A plane wave incident on a grating results in diffracted orders being diffracted at discrete angles with harmonically related spatial frequencies. This is evident in both the reflected and transmitted wave.

Plane wave expansions can be of two types, transverse electric (TE) or transverse magnetic (TM). The special case of a TEM wave may be defined when propagation is defined in the z direction.

Harmonics expansions are simple sine, cosine or exponential functions.

All of the expansions are solutions to Maxwell's equations, or more specifically for the time harmonic case, the Helmholtz equation. For each expansion a number of parameters must be defined, the origin of each expansion, *i.e.*, its location in the x-y plane, the domain in which it is active and in general the order or the expansion. The order of the expansion equates to the how many harmonics are defined. Because of the linearity of the field operators, the field f for each domain can be approximated by f_0, which is the sum of all the expansions defined in that domain. Equation 6.8 describes the field,

$$\begin{aligned} f &= f_0 + \epsilon \\ &= \sum_{j=1}^{J} \sum_{k=1}^{K} a_{jk} f_{jk} + \epsilon \end{aligned} \tag{6.8}$$

where a_{jk} is the coefficient for the jth expansion of order k, f_{jk}, and ϵ is the error between the actual field f and approximated field f_0.

Achieving an optimised solution with high accuracy and minimal computation time depends on judicious choice of expansion type and placement. The initial version of MaX-1 required manual placement of all expansions. Recent versions include functions to semi-automate generation of expansions according to user-defined criteria. While this makes placing the expansions quicker, there is no real inbuilt intelligence as to their placement and the user must have a good understanding of what expansions to place where. Three different algorithms for the expansion placement have been included. The most basic algorithm simply applies a fixed number of expansions at a fixed distance from the boundary. The second algorithm seeks to increase the density of expansions along regions of the boundary with high curvature; the expansions are placed at a distance from the boundary consistent with a range provided by the user. The final algorithm is the most complex and has a number of user definable boxes to constrain the distribution. In the ENFOL simulations a combination of manual and automatic placement (using the second algorithm) has been used.

6.4 THE ENFOL MODEL

The model that formed the basis of the majority of the simulation work is illustrated in Fig. 6.4. It describes a chrome grating of period p suspended in air. Only one period of the grating model (known as a unit cell) needs to be defined in terms of domains, boundaries and expansions. The unit cell then appears to the solver to be repeated infinitely on each side. An *X-periodic* boundary is placed at $y = p$ as one periodic boundary, and the y-axis is defined implicitly as the other periodic boundary. These define the sides of the unit cell. All other boundaries were defined as *Usual*. The grating conductor can also be defined with a surface impedance boundary condition (SIBC), but then information for the fields inside the conductor is unavailable. While this boundary was trialed it was abandoned owing to higher errors in the computation and the fields inside the conductor were also of interest.

Two normally incident plane wave source excitations are considered, with linear polarisation. The transverse electric (TE) excitation is defined with the electric field vector parallel to the grating conductors. The transverse magnetic (TM) excitation is defined with the electric field vector parallel to the grating vector k_g, as defined in Fig. 6.4. It is well-known that solving for the EM fields for highly conducting gratings illuminated with TM polarised light is more difficult than with TE polarised light. Convergence is reported to be slower for RCWA [112] and for FEM [124]. This is owing in part to the discontinuity that exists in the electric field (E_x) in the TM case but not in the TE case. For the MMP models, greater numbers of expansions are required for the TM case compared to the TE case to obtain the same accuracy.

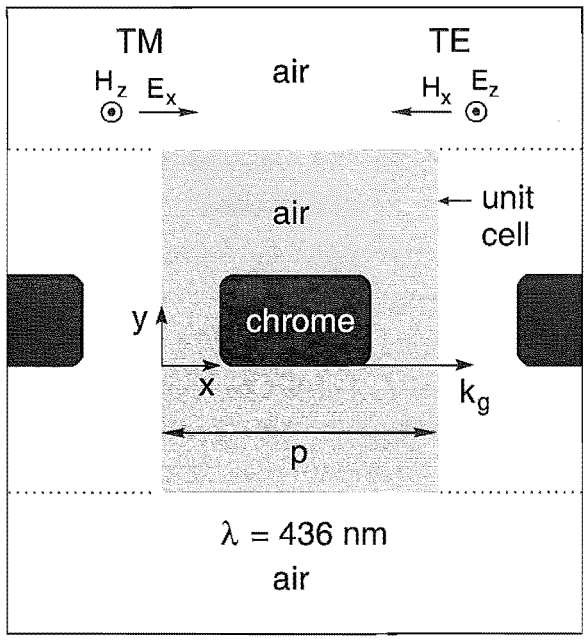


Figure 6.4 Simulation model for suspended grating masks illustrating the TE and TM source polarisation orientations.

The Cr domains were modelled with material properties of $\varepsilon = -13.24 + i14.62$ [94] equating to a skin depth of 17 nm, and $\mu = 1$ at a wavelength of 436 nm. Illumination with 436 nm light was chosen to be consistent with the G-line of a mercury arc lamp, the dominant exposure wavelength for our ENFOL exposures. The air domains were modelled with $\varepsilon = 1$ and $\mu = 1$. Rectangles made from c-polygons define the grating conductors, with the sharp corners replaced by arcs with a 5 nm radius. Sharp corners result in singularities in the EM fields and do not describe the shape of experimentally fabricated conductors. Where the arc meets the straight line, there is a discontinuity in the second derivative, that results in higher errors here, although these can be reduced to tolerable levels by appropriate multipole placement. To remove this second order derivative, the conductor boundary can be approximated by a spline. Using splines means that fewer multipoles are required to obtain the same accuracy, however, to obtain a geometry close to the desired geometry, a large number of spline elements are required. This increases the number of matching points along the boundary, increasing the number of equations to be solved. For this reason spline boundaries were not implemented.

Grating structures were studied as they give a better indication of resolution in comparison to an isolated feature. The visibility of an isolated feature only gives half the story, as it is easier to resolve than two features close together with the deleterious broadening owing to diffraction. Simulating two isolated features would be valid, however it is actually a larger problem to solve than a grating as MMP is able to take advantage of the periodicity of the problem.

6.4.1 Placement of Multipoles

Three types of expansion were used. A plane wave expansion provided the source excitation, and Rayleigh expansions were used to describe the reflected and transmitted fields at the entry and exit of the unit cell respectively. The rest of the fields were described by multipole expansions, the most commonly used expansion type. The following are some guidelines to placing expansions, they are derived from personal trial and error, and from comments made by Christian Hafner [120], [125]. At least one multipole is required for each side of a boundary and usually more unless the geometry is particularly simple. In general more accurate solutions can be obtained with larger numbers of multipoles, assuming they are placed sensibly, however, this comes at a cost of computation time. It is useful to consider each multipole's area of greatest influence for multipoles placed sequentially along a straight boundary. A multipole's area of greatest influence encompasses a triangular area from the multipole a distance d away from the boundary and extending approximately $2d$ along the boundary as illustrated in Fig. 6.3. Placing poles closer than $2d$ apart is not recommended as dependencies between multipoles can occur, leading to numerical instabilities and large errors. Placing multipoles too far apart means that the field can be under-constrained, and unable to satisfy the boundary conditions sufficiently, resulting in large errors. For arc-shaped segments of the boundary, multipoles can generally be placed in the middle of the arc. This is acceptable when the field is slowly varying, however in the case of rectangular gratings illuminated by a TM polarised wave, a discontinuity in the field at the grating corners means that more multipoles are required to more accurately describe the field here. This introduces another rule of thumb, regions with a high degree of curvature or where the field has a high degree of spatial variance require more poles.

Figures of merit as to the accuracy of solutions for a simulation model are provided following the determination of the expansion coefficients. A maximum and average absolute error are reported which tell the user the greatest, and the average mismatch in the fields according to the boundary conditions at the matching points. A maximum and relative error are also provided which are the errors normalised to the size of the field at the matching point. A graphical view of the error at the matching points is also provided that indicates where the placement of poles can be improved; again, a trial and error approach. While these errors provide useful information for fine-tuning a model, they can not be wholly relied upon. It is possible to over-constrain the model by using large numbers of multipoles. If they are very close to the boundaries, their area of influence can be quite small, and while the boundary conditions may be satisfied the field may be inadequately modelled in the spaces between the boundaries. This was noted by the author when large period gratings were being simulated and the distance between the conductors and the bottom of the unit cell was large compared to the

distance between the poles and the boundary they acted on.

6.4.2 Simulation Outputs

A number of outputs are calculated by MaX-1, including the electric field and magnetic field vectors, the Poynting vector representing the power flow, and various field densities, voltage densities and current densities. For photolithography purposes, it is the square of the electric field ($|E|^2$) that is of most interest as it has the strongest influence on an exposure. Photoresist is a lossy dielectric and it is the complex term in the photoresist's permittivity that determines the energy absorbed during the exposure. Part of the absorbed energy is responsible for effecting a chemical change in the photoresist. That $|E|^2$ is the term of interest, as opposed to the Poynting vector, is experimentally borne out by the formation of standing wave patterns visible in photoresist following a monochromatic exposure and development with reflective substrates [77]. A standing wave pattern is not evident in the Poynting vector.

The electric field is extracted from MaX-1 and results are presented in terms of a normalised intensity $|E|^2/|E_{in}|^2$ where E_{in} is the time averaged incident electric field. While this gives a good indication of the intensity distribution in the photoresist it does not tell the whole story. In reality the resultant resist profile is a more complex problem. Photoresists have a non-linear response to intensity and during an exposure "bleaching" occurs, a process whereby the photoresist becomes less absorbing. The subsequent development is a reaction limited process, the dissolution of the exposed resist only occurring at the photoresist/developer interface. While these issues clearly have an effect on resist profile, this work only considers the intensity distribution.

6.5 SIMULATIONS RESULTS FOR SUSPENDED GRATINGS

This section analyses the diffracted fields from a lossy chrome grating suspended in air. Investigation of this simplified system highlights the major trends without the complications of reflections due to multi-layers with mismatched refractive indices. This allows the more fundamental limitations of the system to be evaluated. The simulation results presented in this section were performed using models of the form shown in Fig. 6.4. Gratings with 1:1 line and space ratios for periods p from 200 nm down to 20 nm were modelled in 40 nm thick chrome and illuminated with a 436 nm linearly polarised source.

Figure 6.5 shows the 2-D normalised intensity $|E|^2/|E_{in}|^2$, where E_{in} is the incident electric field, for an evanescent exposure of a 140 nm-period Cr grating suspended in air. In Fig. 6.5(a) the resultant intensities from TM incident light are illustrated. In the first 50 nm or so below the grating a high contrast image of the grating is apparent. Enhanced intensities, greater than the incident intensity, are evident in the apertures

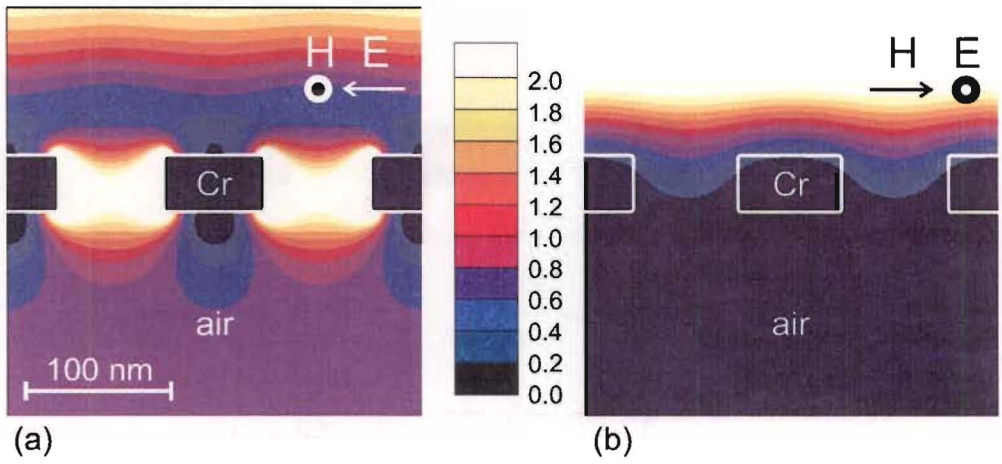


Figure 6.5 Normalised intensity distribution for a 140 nm-period, 40 nm-thick Cr grating in air. The grating is illuminated from above by 436 nm linearly polarised light with (a) TM polarisation and (b) TE polarisation (electric field E and magnetic field H directions are indicated). Contour plots of the normalised electric field intensity are shown $|E|^2/|E_{in}|^2$, where E_{in} is the incident electric field. The scale varies linearly from 0 (black) to 2.0 (white) in 10 linear steps.

below the exit plane of the grating and a shadowed region exists below the Cr absorbers. Extending beyond this region, uniform intensity is observed. A halo of enhanced fields surround the Cr side-walls. These encroach slightly around the corners to the top and bottom surfaces, expanding the exposed region close to the grating. The intensity map for the grating illuminated with TE light (Fig. 6.5(b)) shows a quite different picture. For this polarisation there is very little transmitted light. The grating acts like a wire-grid polariser, transmitting the TM light, but very little TE light. No edge enhancements are evident in the TE polarisation and very little image contrast is visible beneath the grating. Areas of high intensities are evident in regions above the grating for both polarisations. These are standing waves resulting from the interference of the incident and reflected light.

Line-plots taken from the simulations presented in Fig. 6.5 are illustrated in Fig. 6.6. They show the normalised intensities for TM and TE exposures at different depths, within the grating and below. Figure 6.6 quantitatively demonstrates the dominance of the TM exposure (solid line) over the TE exposure (dotted line) at this grating period. While some contrast is evident in the TE exposure within the grating, it has all but disappeared by the exit aperture of the grating.

The TM exposure clearly shows how the field intensity changes as a function of depth behind the mask. Good confinement of the light is obtained at a cross-section through the centre of the grating metal Fig. 6.6(a). At the exit plane, in Fig. 6.6(b) strong field enhancements are exhibited at the grating edges. At 10 nm and 20 nm below the mask (Fig. 6.6(c),(d)), the field enhancements are dampened considerably while significant intensity beneath the aperture is still maintained. By 50 nm below the mask

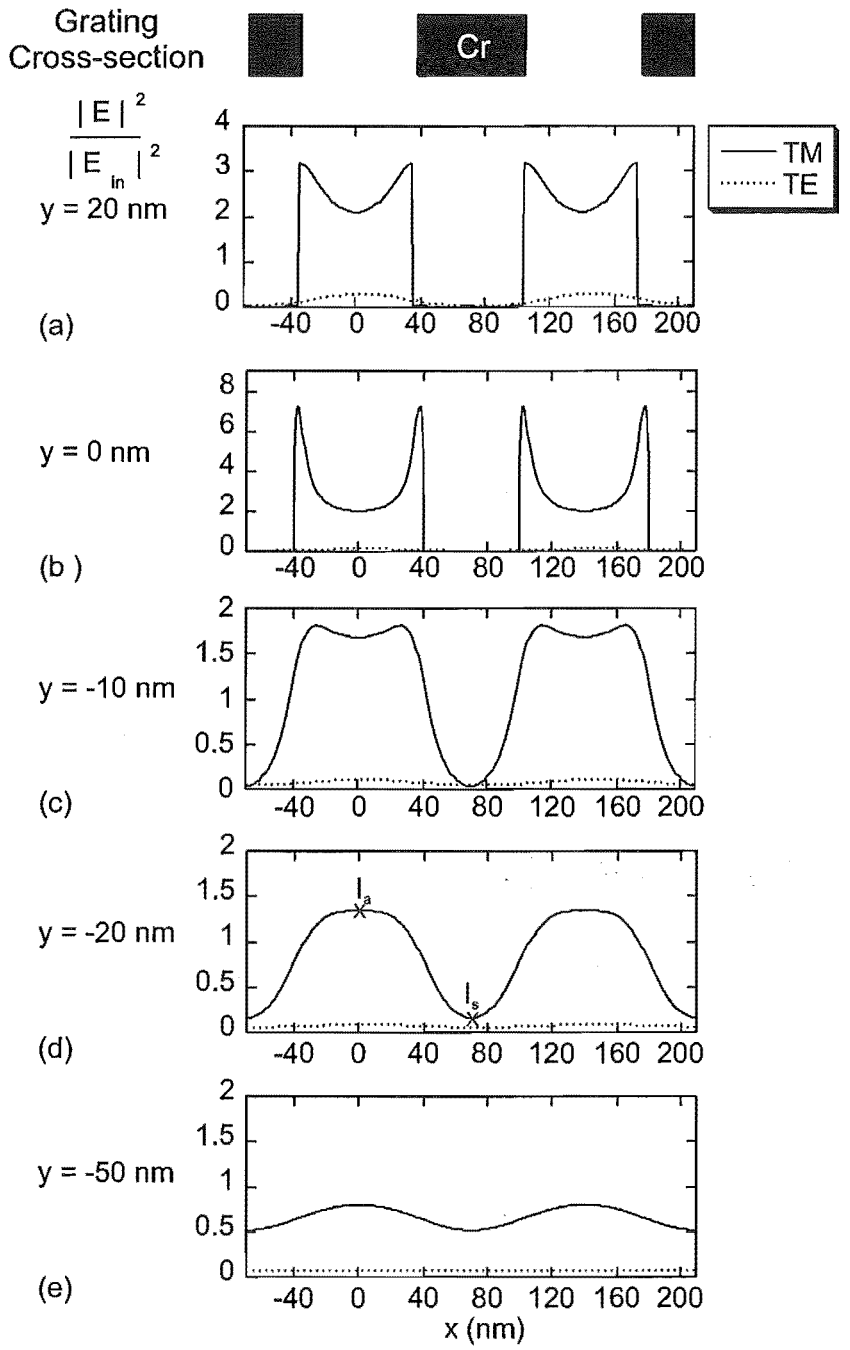


Figure 6.6 Normalized intensity line plots for the simulations described for Fig. 6.5, intensities resulting from TM illumination (solid line), and the intensities resulting from the TE illumination (dotted line). The line plots are taken at (a) the centre of the grating conductors (y = 20 nm) (b) y = 0 (the exit plane of the grating), then at (c) 10 nm, (d) 20 nm and (e) 50 nm below the exit plane.

(Fig. 6.6(e)) only a small amplitude modulated signal remains. These figures illustrate two characteristics of an ENFOL exposure — the decaying amplitude of the intensity modulation as we move further below the grating, and strong edge enhancements that are evident close to the exit aperture of the grating. The former will be discussed in Section 6.5.1 and the latter in Section 6.5.3.

The simulation results presented in Fig. 6.5 and Fig. 6.6 correspond to an evanescent exposure. An evanescent exposure is defined as an exposure where all the transmitted diffracted orders (excepting the zeroth order which is effectively undiffracted) are evanescent. For incident illumination normal to the grating an evanescent exposure occurs when $p < \lambda/n$, where λ/n is the effective wavelength, as detailed in Section 2.6. An evanescent exposure is typified by enhanced image contrast immediately below the grating that decays away exponentially, typically within $p/2$. Thereafter only a uniform intensity exists from the zeroth transmitted order. While the zeroth diffracted order is propagating, it does not provide contrast for photolithography since it is unmodulated in the x direction. The remainder of this chapter investigates the important features of evanescent exposures via simulation. Simulations of large period structures and the transition to evanescent exposures are presented in Chapter 7.

6.5.1 Effect of Decreasing the Grating Period

Simulation results for Cr gratings suspended in air reveal some interesting trends as the grating period is reduced from 200 nm to 20 nm, with notable differences between the two polarisations. In Fig. 6.7 the normalised intensity profiles (normalised with respect to the incident intensity) are shown in the exit plane of the grating (y = 0 nm) for the different periods. As observed in the line-plots in Fig. 6.6, the TM polarisation (solid) leads to greater intensities behind the mask compared to the TE polarisation (dashed). Note that the TE results have been scaled by a factor to give the same intensity as the TM results at the centre of the grating aperture ($x = 0$). The scaling factors are indicated on each of the graphs. The TE intensities fall sharply away with decreasing period, which is to be expected as these gratings act as polarisers, with the polarisation efficiency improving as the period of the grating is reduced.

Notably the 20 nm periods are clearly resolved at the exit aperture as shown in Fig. 6.7(d) for the TM polarisation. The intensity at the apertures is enhanced relative to twice the incident field and a strong null is evident underneath the absorber. This indicates the enticing prospect of resolutions better than $\lambda/20$.

The most noticeable feature of Fig. 6.7 is the large edge enhancements evident at the grating edges for the TM polarisation, observed previously for the 140 nm period grating. They consist of high spatial frequency components that decay rapidly. The field enhancements weaken as the period is reduced, and by the 20 nm-period grating in Fig. 6.7(d), all sign of the edge enhancements have disappeared. A weakened edge

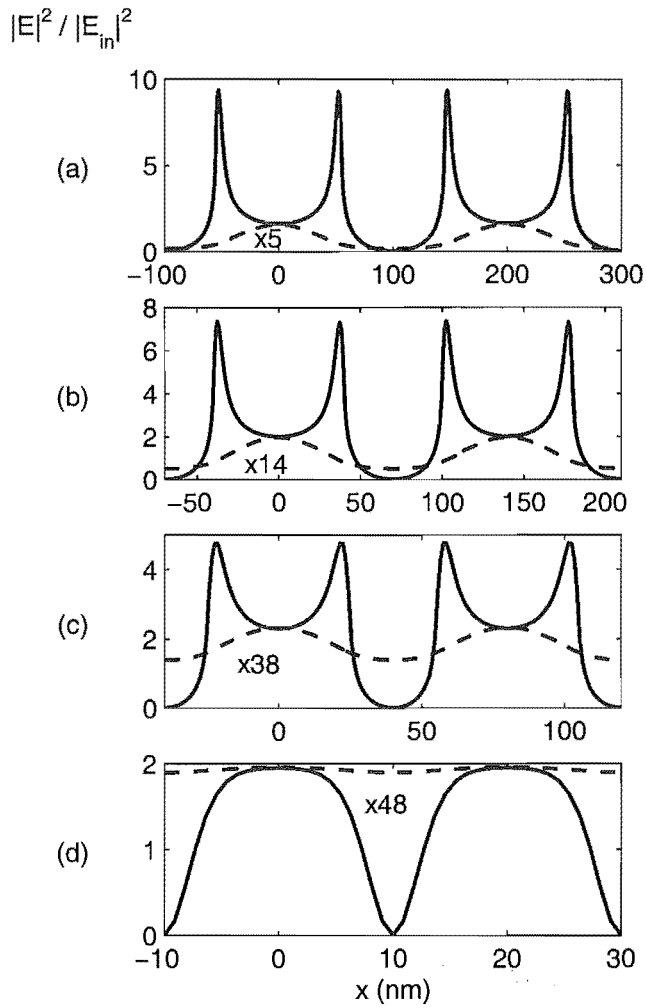


Figure 6.7 Normalised intensity at the exit plane of gratings with periods of (a) 200 nm, (b) 140 nm, (c) 80 nm, (d) 20 nm. TM and TE excitation are the solid and dashed lines respectively. Note that the fields resulting from the TE excitation have been scaled (scale factor indicated on graphs).

enhancement for the 20 nm-period grating is present at $\approx$ 2 nm above the exit aperture, close to the effective corner of the 5 nm radiused conductor. However, by the exit aperture the enhancement has already died away.

The steepness of the edge enhancements suggest that if suitable exposure conditions were selected, a period halving may be obtained. This has not been observed experimentally, however the small dimensions of the enhancement may be beyond the resolution capabilities of our Shipley G-line photoresist. The edge enhancements are laterally quite confined if an optimal threshold is chosen near the steepest part of the curve. The enhancement is also bound closely to the exit aperture of the grating, visible only in the first 10–20 nm.

Figure 6.8 illustrates the range of the evanescent near field as a function of the different grating periods in terms of the image contrast or visibility V. The image

contrast is defined as

$$V = \frac{I_a - I_s}{I_a + I_s} \tag{6.9}$$

where I_a is the intensity at the centre of the aperture and I_s is the intensity at the centre of the shadow; these are identified in (d) of Fig. 6.6. The TM polarisation

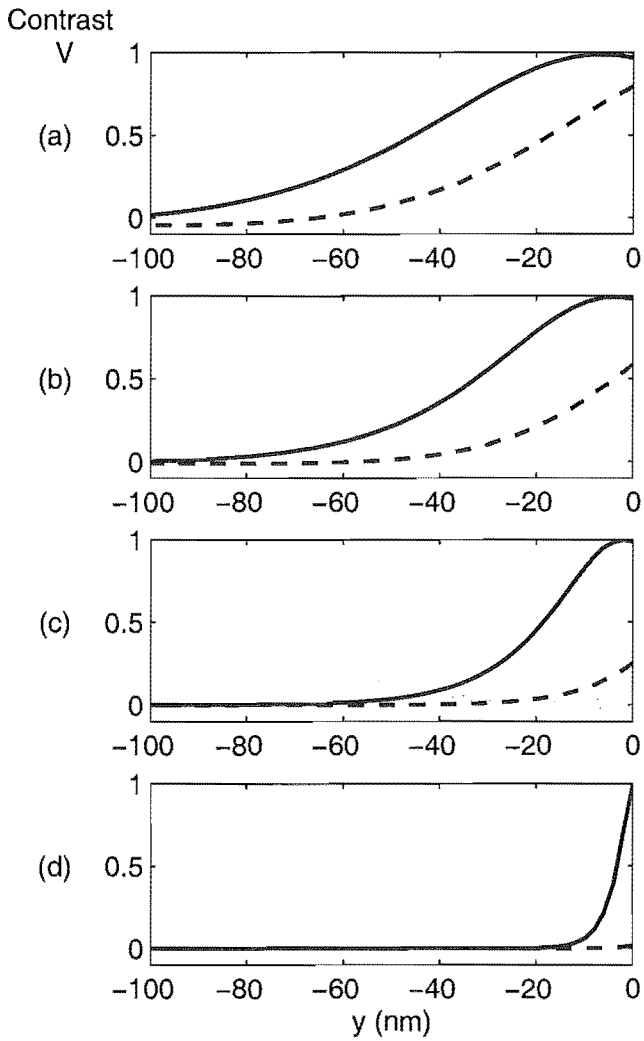


Figure 6.8 Contrast as a function of depth from the mask for gratings of various grating periods (a) 200 nm, (b) 140 nm, (c) 80 nm, (d) 20 nm. TM and TE excitation are the solid and dashed lines respectively.

contrast (solid) is larger than that for the TE polarisation (dashed) for all grating periods and at all distances from the grating. For the 80nm- and 20nm-period gratings, the contrast for the TE mode is less than 0.3 even at the exit plane of the grating. In comparison no such reduction in contrast is seen for the TM polarisation, however there is a noticeable shortening of the high contrast region, which is confined to increasingly shorter distances from the grating at decreasing periods. For a contact lithography

application this equates to requiring thinner resists to ensure development right through the resist as the period is decreased.

High contrast for lithography is essential, but viability for an evanescent near-field exposure technique also depends on having sufficient power to avoid unreasonably long exposure times. The normalised intensity in the middle of the aperture is plotted as a function of distance from the back of the mask in Fig. 6.9. Again the polarising effect

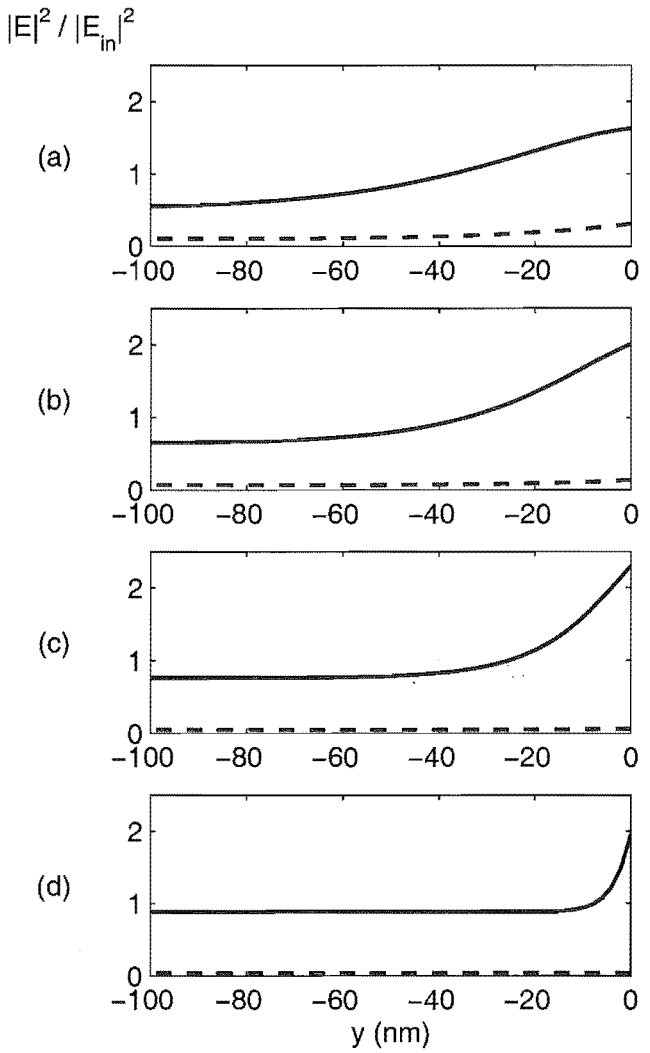


Figure 6.9 Intensity at the centre of the aperture as a function of distance from the grating for periods of (a) 200 nm, (b) 140 nm, (c) 80 nm, (d) 20 nm. TM and TE excitation are the solid and dashed lines respectively.

of the grating mask is evident, with ample intensity in the TM polarisation and the TE polarisation displaying low intensity at all depths and grating periods. The remainder of this chapter will focus predominately on results using TM illumination.

Worthy of note for the TM polarisation is the field enhancement in the centre of the aperture at distances close to the mask. The enhancement decays with distance

to the far-field intensity value (resulting from the zeroth transmitted order). Here the grating contrast is negligible, the intensity profile flat across the grating. The far-field intensity for the TM polarisation increases as the grating period is reduced, approaching the incident intensity, the grating performing closer to an ideal polariser. The increased transmission of smaller period gratings can be understood qualitatively by considering the response of the grating conductors to the incident electric field. An incident electric field causes charge in the conductors to separate, which in turn acts to cancel the electric field at the air/conductor boundary. The charge in smaller grating conductors (with respect to the wavelength) is able to separate faster, to more effectively cancel the electric field at the air/conductor boundary. The electric field is concentrated across the grating apertures and the energy is guided through the apertures that behave like parallel-plate waveguides. If incomplete cancellation occurs at the conductors, as is the case with larger grating conductors, less of the electric field is across the apertures. Less energy is therefore guided through the grating and there is increased reflection at the conductors. The parallel-plate waveguides do not transmit the TE mode as the electric field is tangential to the long dimension of the plates.

6.5.1.1 The Effect of Grating Period on Depth of Field

The trend of decreasing contrast with depth below the grating is clearly important to quantify for lithography. The depth of field (DOF) partially embodies this information practically by defining the depth at which "sufficient" contrast is available for an exposure. This is not a fixed value however, it varies with different resist chemistries and different exposure conditions. For the purposes of this thesis we will define DOF_V as the depth below the grating at which the contrast V falls below a specified value.

Figure 6.10 plots the depth of field as a function of the grating period, extracted for various values of V. These results show that the DOF_V decreases linearly with period. The depth of field is greater for smaller values of V, as might be expected. This implies that by using a high contrast resist, thicker resists can be used which can ease constraints for pattern transfer. From the graph, gratings with a 300 nm period, at the lowest contrast factor ($V = 0.3$), indicate depths of field of less than 100 nm. This is a conservative estimate however, as experimentally the exposure is into a photoresist with a refractive index greater than one.

The linear nature of the relationship between the DOF and period can be qualitatively understood for an evanescent exposure by considering y_m^I, the depth at which each diffracted component decays to $1/e$ of its initial intensity [126],

$$y_m^I = \frac{1}{4\pi\sqrt{\frac{m^2}{p^2} - \frac{n^2}{\lambda^2}}} \tag{6.10}$$

where m is the diffracted order, and n the refractive index of the material surrounding

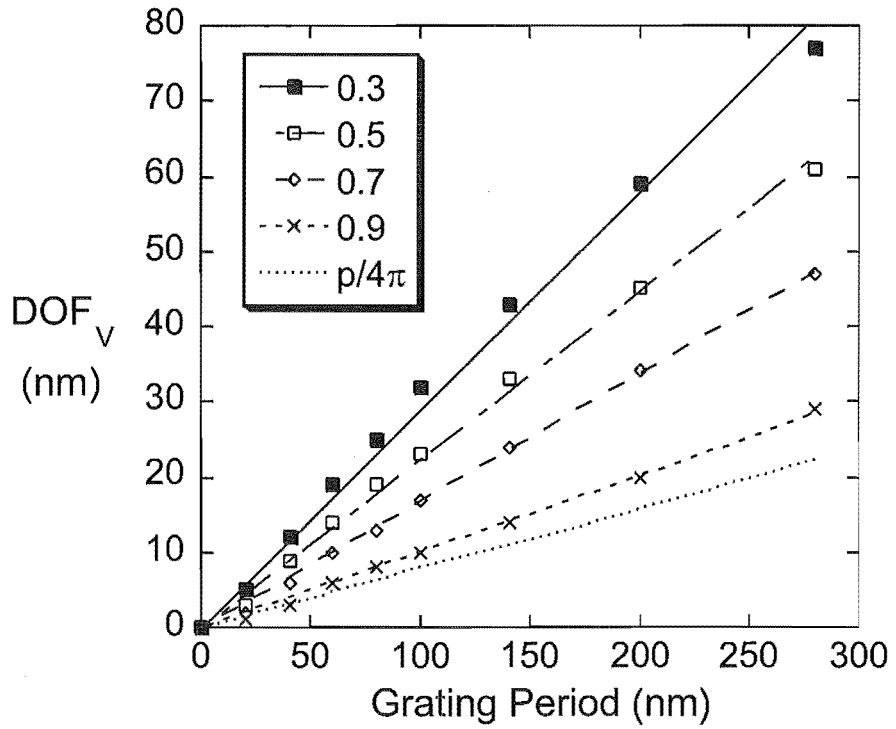


Figure 6.10 Depth of Field (DOF_V) versus period p for simulated gratings plotted for V values of 0.3, 0.5, 0.7, 0.9. The depth corresponding to 1/e of the intensity of transmitted light of the first order ($p/4\pi$) is also plotted.

the grating. From Equ. 6.10 it is evident that the higher the evanescent diffracted order, the faster its intensity decays. This has an impact on the exposure at the depth defined by DOF_V. For high values of V there exists a significant contribution from higher order ($m > 1$) evanescent components while at low V values there is generally only a significant contribution from the $\pm$ 1 orders to provide contrast.

For gratings with deep sub-wavelength periods $p \ll \lambda/n$, Equ. 6.10 approximates the linear relation $y_m^I \simeq p/4\pi m$. Note that this is independent of λ and n, so we expect the resolution of ENFOL in this regime to be independent of these parameters. This opens the prospect of choosing an exposure wavelength for optimal resist performance, by-passing the usually difficult task of designing resist performance for an optimal exposure wavelength.

The line equating to $y_{\pm 1}^I = p/4\pi$ is plotted in Fig. 6.10. The intensity from the first diffracted order is the dominant component providing contrast for an exposure, particularly as we move further away from the grating where the higher order evanescent components have disappeared. In this region the intensity consists of the zeroth-diffracted order modulated by the contribution from the first evanescent diffracted orders, this is shown for example in Fig. 6.6(d). The $p/4\pi$ line falls below the V = 0.9 on the graph, which means that the intensity for the $\pm$1 diffracted orders has al-

ready decayed to 1/e of its initial value at depths corresponding to high contrast. This suggests that enhancement of the ± 1 diffracted orders is partially responsible for providing high contrast. The relative magnitudes of the zeroth and first-order components also have an important effect on the depth of field.

The choice of resist thickness is crucial for a successful exposure, firstly to obtain sufficient contrast and secondly to ensure adequate process latitude. It is clear from Fig. 6.10 that the resist thickness should be chosen according to the smallest feature period to be patterned. To improve process latitude, operation at a high V is preferable which equates to reducing the resist thickness further for a given resist system. Operating at low V values increases the likelihood of exposure variation, as the steepness of the intensity curves is reduced, making the threshold of exposed or unexposed resist less well-defined.

For the smallest period plotted in Fig. 6.10 (20 nm), a depth of field of approximately 3 nm is indicated, at $V = 0.5$. While the depth of field is not large, such resolution may be experimentally achievable using new generation resist chemistries such as surface layer imaging resists [127] or self-assembled monolayer resists [128].

6.5.1.2 The Effect of Skin Depth on Resolution

One might expect the resolution of ENFOL to be limited by the skin depth of the mask metal, however this is not the case. Figure 6.11 illustrates the field intensity for a 20 nm-period grating as a function of depth. The grating conductors are modelled with realistic material properties and the calculated skin depth for the metal at 436 nm is 17 nm. This is thicker than the widths of the grating conductors, which are only 10 nm. In Fig. 6.11 (a) we can see that the intensity profiles have an extremely steep gradient at the edges of the grating. This is significant as it implies that there is little field penetration in the lateral direction for the TM polarisation, despite the large skin depth with respect to the conductor width. Intuitively this does not seem reasonable but it can be understood if the surface charge that exists at the grating sidewall is taken into consideration. The impinging electric field on the grating sidewalls for the TM polarisation is predominately in the x direction for a good conductor. By using a quasi-static argument and Gauss' Law the net charge on the surface of the grating accounts for the discontinuity of the electric field normal to the grating. Inside the grating the electric field drops to a value close to zero, then decays exponentially. This results in good lateral field confinement at the exit aperture of the grating Fig. 6.11(b). The same argument does not hold for the top surface of the grating metal since here the electric field is tangential to the grating surface and to satisfy the boundary conditions, the tangential components on either side of the boundary must be equal. There is no discontinuity in the electric field here and inside the grating the electric field decays exponentially resulting in significant field penetration for a 40 nm thick grating with a

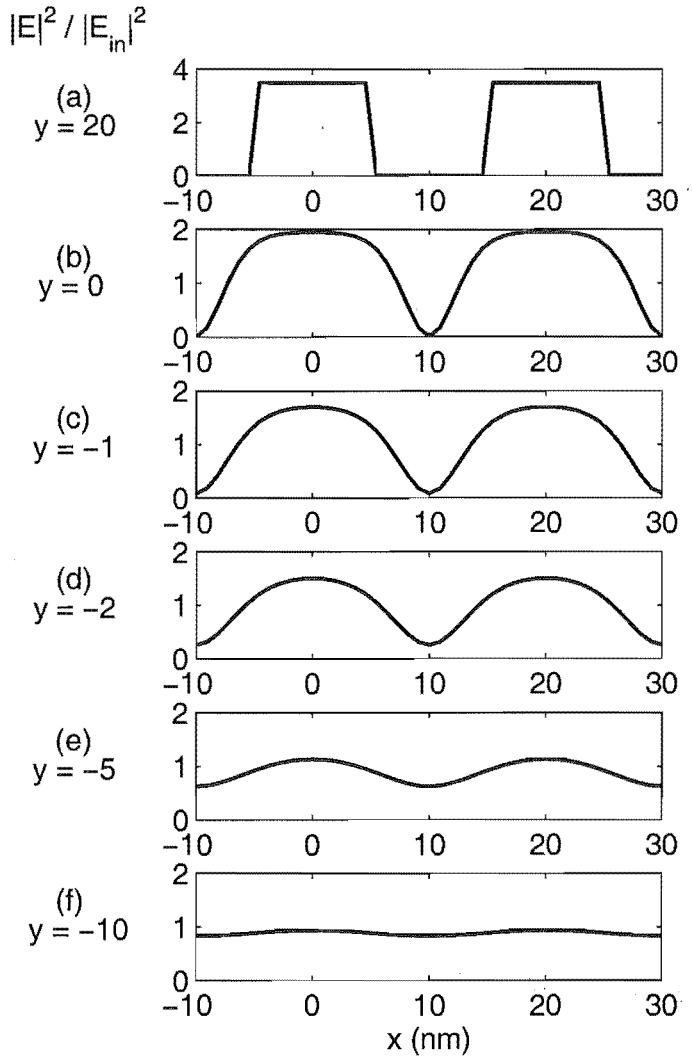


Figure 6.11 Normalised intensity profiles for a 20 nm period grating with TM excitation at (a) the centre of the grating, (b) the exit plane of the grating, (c) 1 nm below the grating, (d) 2 nm below the grating, (e) 5 nm below the grating and (f) 10 nm below the grating.

skin depth of 17.1 nm. The reduced sensitivity to skin depth of small period structures in the lateral direction is potentially very useful for high resolution lithography, however the thickness of the conductor remains important relative to the skin depth if reasonable shadowing is to be obtained.

6.5.2 Effect of Grating Duty Cycle

The effect of the grating duty cycle on an evanescent exposure was investigated. Note that duty cycle is defined here as the percentage of conductor width relative to the period. As has been shown in the previous sections for 50% duty cycle gratings, with TM incident light there is adequate exposure intensity to expose a resist at a restricted depth below the grating. Changing the duty cycle of a polariser has a strong influence

on the intensity of the preferentially transmitted polarisation, which suggests it may be an important parameter for an ENFOL exposure.

A range of duty cycles from 10 to 90% were investigated for a 140nm period grating, with otherwise the same model parameters as previously described. Figure 6.12 shows normalised intensity field plots for a 30% duty cycle grating (Fig. 6.12(a)), and a 70% duty cycle grating (Fig. 6.12(b)). Comparing the two plots, the 30% duty cycle grating has a much higher background intensity (equating to a higher zeroth order transmission coefficient T_0) than the 70% grating. It is the 70% duty cycle grating that appears to have the higher contrast immediately below the exit aperture. The electric fields in the aperture are more intense, and the shadowed regions less intense suggesting the higher duty cycle has more desirable properties for an ENFOL exposure.

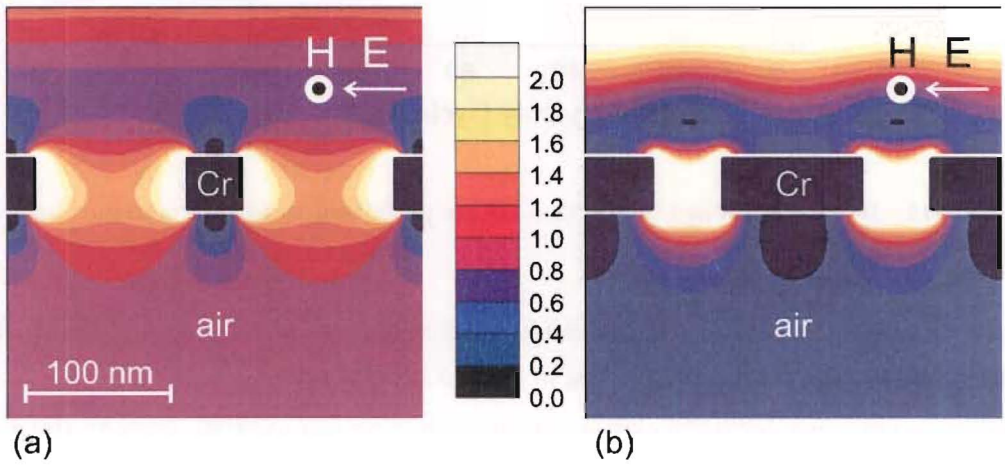


Figure 6.12 Normalised intensity distribution for a 140 nm period, 40 nm thick Cr grating in air for two different duty cycles. In (a) a 30:70% duty cycle is presented with 42 nm wide Cr absorbers and 98 nm apertures. In (b) a 70:30% duty cycle grating is presented with 98 nm wide Cr absorbers and 42 nm apertures. The grating is illuminated from above by 436 nm TM polarized light (electric field E and magnetic field H directions are indicated). Contour plots of the normalised electric field intensity are shown $|E|^2/|E_{in}|^2$, where E_{in} is the incident electric field. The scale varies linearly from 0 (black) to 2.0 (white) in 10 linear steps.

Figure 6.13 plots the depth of field for a V equal to 0.5 ($DOF_{0.5}$) as a function of duty cycle on the left axis, and the maximum normalised intensity in the aperture at this depth is plotted on the right-hand axis. As the duty cycle increases, a steadily improving depth of field is observed. This improvement is largely a result of the improved ratio between the higher order diffracted components to T_0, as T_0 reduces with increasing duty cycle. This improvement comes, however, at the cost of exposure intensity as indicated on the right-hand axis of Fig. 6.13, longer exposure times being required for higher duty cycle gratings. This impacts not only on throughput, but more importantly the strong duty cycle dependence could complicate the exposure of masks with mixed grating periods. One way of mitigating these effects is to operate with resist thicknesses consistent with the grating with the minimum depth of field. This is

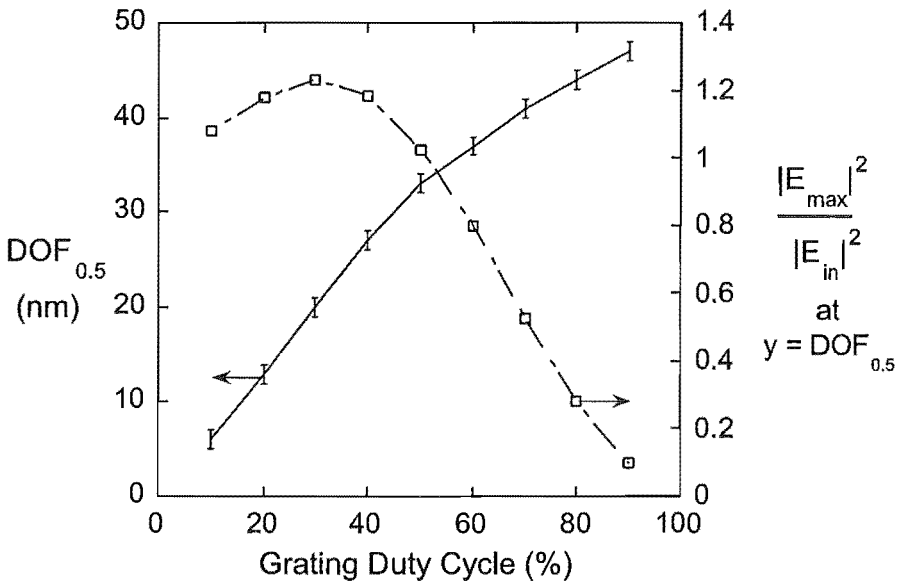


Figure 6.13 Depth of Field at $V = 0.5$ for a 140 nm period grating and the maximum intensity at this depth versus duty cycle.

similar in many respects to the necessary requirement of considering a resist thickness appropriate for the minimum grating period to be exposed.

These exposure dependencies on physical grating parameters, such as duty cycle and period, are a feature of sub-wavelength structures. While this work only investigates this effect for grating-like structures, similar feature size dependencies are expected for other feature geometries, although this must be the subject of future work.

6.5.3 Effect of Grating Conductor Radius of Curvature

Close to the exit aperture of the grating, high intensity enhancements occur at the grating conductor edges, as can be seen in Fig. 6.6(a). These enhancements are due to the sharp discontinuity at the conductor corners which encourages charge concentration in a manner analogous to a static electromagnetic problem. The static analogy becomes more reasonable as the grating period becomes much smaller than the wavelength. These enhancements are only evident with TM illumination as it is only in the TM case that the electric field discontinuity exists.

Simulations were carried out for a 140 nm-period grating, with the same conditions as those described for Fig. 6.5, but with the radius of curvature of the conductor corners varied from 1 to 10 nm in steps of 1 nm. An increase in amplitude of the edge enhancements is evident as the radius is decreased, consistent with the static analogy. This is shown in Fig. 6.14 where we see the peak intensity at the exit aperture for a

grating with a 1 nm radius is 25 times the incident intensity and six times greater than for a 10 nm radius. The relationship between the radius and peak intensity follows a power-law decay, as may be expected from a simple electrostatic analogy. A power-law curve can be fitted for the intensity, with the radius as the dependent variable and an exponent of -0.779 obtained with high correlation to the data. This is between the static field analogies of the intensity arising from a cylindrical geometry, $I_{cylinder}$, where $I_{cylinder} \propto 1/r^2$ and the intensity due to two conducting plates meeting at a 90° angle $I_{90°}$, where $I_{90°} \propto 1/r^{2/3}$ [129].

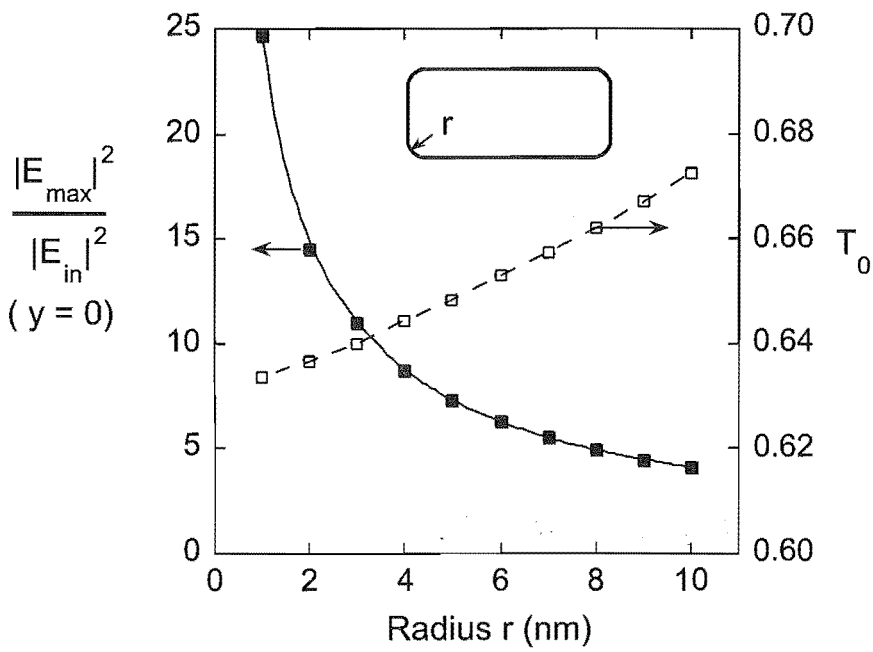


Figure 6.14 The peak intensity at the grating exit aperture is plotted on the left axis, while on the right axis, the zeroth order intensity coefficient T_0 versus the grating conductor radii of curvature r. The simulation model is the same as that of Fig. 6.5.

The edge enhancements are relatively short-range. In 20 nm the peaks become less intense in amplitude than at the centre of the aperture, and their decay is independent of the radii of curvature. In the far-field an increase in T_0 is evident at the expense of other diffracted orders as the radii are increased, as shown in Fig. 6.14. This acts to reduce the contrast and, over the range of radii simulated, the zeroth transmitted order increases by 4%. This suggests that for reliable lithography, maintaining as small a radius as possible is desirable. While small radii equate to more intense field enhancements, it is not anticipated that these should be a problem when operating with resist thicknesses greater than the range of the field enhancements. However, this assumes that the photoresists' response to a strong localised field results in a similarly localised bleaching.

6.5.4 Effect of Grating Conductor Conductivity

The effect of the grating material conductivity on an evanescent exposure was investigated by comparing the resultant fields for different materials. Metals with high conductivities, such as chrome (Cr) and gold (Au) were compared. Tungsten (W), a metal which is not as conductive, was also investigated as it is a possible alternative material for the absorbers on ENFOL masks. Tungsten has the desirable property of being able to be dry etched with gases available with our RIE system. This would allow the formation of mask patterns using an etch-back technique, avoiding the mask profile issues resulting with lift-off, discussed in Section 4.5. Lastly, silicon (Si), a dielectric with a high permittivity is included for comparison.

A table of their relevant material properties, such as their complex refractive indices, relative permittivities and skin depths are presented in Table 6.2. The materials' complex refractive indices were obtained from published tables [94] at an energy of 2.8 eV (or as close as possible) which equates to a λ of $\sim$ 436 nm. The complex permittivities that were used in the simulations were derived according to Equ. 6.7.

Table 6.2 Optical parameters for different materials at 436 nm (2.84 eV).

Material	n'	n''	ε_{mat}	δ (nm)
Au	1.46	1.77	-1.001 + i5.168	39.8
Cr	1.80	4.06	-13.24 + i14.616	17.1
Si	4.75	0.16	22.564 + i1.549	425.5
W	3.31	2.49	4.756 + i16.484	27.8

The skin depth δ, is defined as the depth into the metal where the field has decayed to $1/e$ of its value at the surface, was calculated from Equ. 6.11,

$$\delta = \frac{c}{2\pi f \Im(\sqrt{\varepsilon_{mat}})} \tag{6.11}$$

where c is the speed of light in a vacuum, f the frequency of the radiation, and $\Im(\sqrt{\varepsilon_{mat}})$ represents the imaginary part of the square root of the relative permittivity of the material, ε_{mat}.

In general, the smaller the skin depth, the better the conductor. Table 6.2 shows that Cr has the smallest skin depth, 17 nm at λ = 436 nm. This is less than half the thickness of the grating conductor for these simulations. Tungsten also has a small skin depth, however at this wavelength it exhibits strong dielectric behaviour. Metals are generally defined as having a negative real component of the permittivity [130] and W has a positive real component as $n' > n''$. Tungsten is a material whose properties lie on the continuum between metallic and dielectric behaviour, the dominant behaviour being strongly wavelength dependent. Note that describing W and Si as having a skin depth is not strictly correct considering their dielectric nature, however it does

give an indication of the field penetration. W's small skin depth, only slightly greater than Cr and smaller than Au, is due in part to its strong absorbance which is directly proportional to n''. The absorption coefficient α is defined as

$$\alpha = \frac{4\pi n''}{\lambda_0} \tag{6.12}$$

where λ_0 is the free space wavelength.

Simulations of a 140 nm period grating with TM and TE polarised light at 436 nm were performed with 40nm thick gratings of Cr, Au, W, and Si. The resultant intensity maps for the four materials show some interesting results. From the exposure results for TM light shown in Fig. 6.15 we see that all four materials have similar intensity patterns with a high field region in the apertures and some shadowing behind the grating absorbers. Note that each of the materials, even the dielectrics, show a strong field enhancement at the grating exit aperture, as illustrated in the line plot in Fig. 6.15(b). Cr appears to have the highest contrast, with the strongest field null immediately below the grating fingers, and the highest intensity in the apertures. Si demonstrates the lowest contrast with higher intensity fields present more uniformly below the grating. Surprisingly W appears to have higher contrast than Au, although Au has an overall lower background intensity.

The resultant exposures with the TE polarisation illustrated in Fig. 6.16(a) show more noticeable differences between the different materials. Again Cr appears to have the best contrast with a reasonably high amount of transmission through the Au grating as the grating thickness is approximately equal to the skin depth of Au. The field map for W shows very little transmission, while Si shows an inverted grating image, with high intensities behind the apertures, and low intensities in the apertures. The high ε of Si acts to concentrate the field in the Si regions, the same mechanism used for the LCM exposures described in Section 1.3.4.3. There is a field null at the top surface of the Si where there is destructive interference as the incident plane wave travels from an area of low to high permittivity. At the exit aperture, the opposite occurs with the primary incident wave travelling from a high to a low permittivity creating an intensity maximum. This behaviour is echoed in the W as shown in the line plots in Fig. 6.16(b) taken at the exit aperture of the gratings. Due to the larger losses in the W, the intensity maximum is not as high.

Overall, while W is not strongly conductive, it may be a suitable candidate for the grating material, better in fact than Au. It shows greater contrast than Au in the TM polarisation and weaker fields behind the conductor in the TE polarisation. For Au to be a suitable candidate, a thicker layer would be required. The superior performance of Cr compared to the alternative materials illustrate why Cr has been the standard for photomasks at these wavelengths.

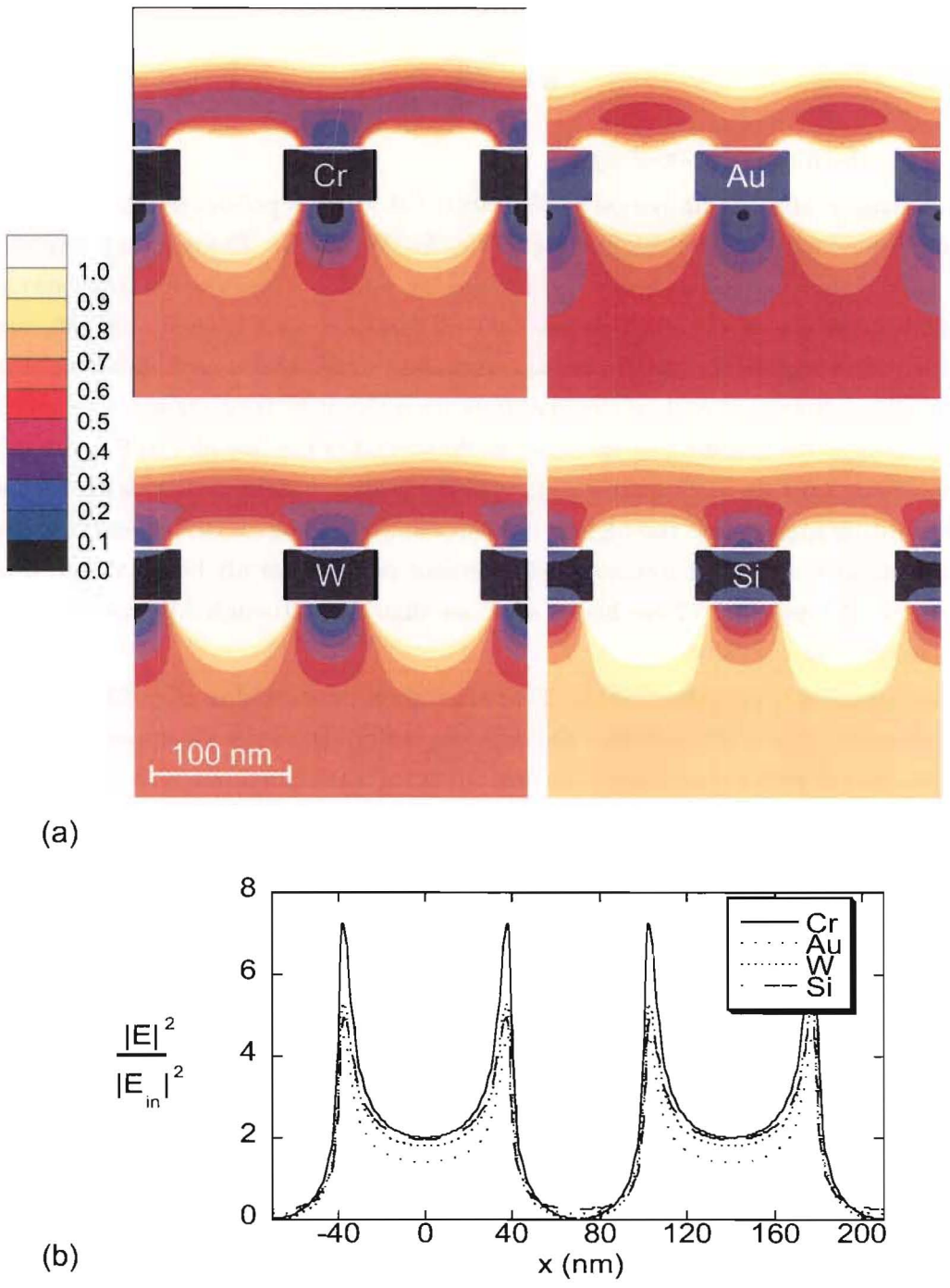


Figure 6.15 Normalised electric field intensity for a 140 nm-period grating illuminated with TM polarised light, for Cr, Au, W, and Si grating materials. (a) field maps with contour lines varying linearly from 0 (black) to 1.0 (white) in 10 linear steps. (b) Normalised intensity at the exit aperture of the gratings.

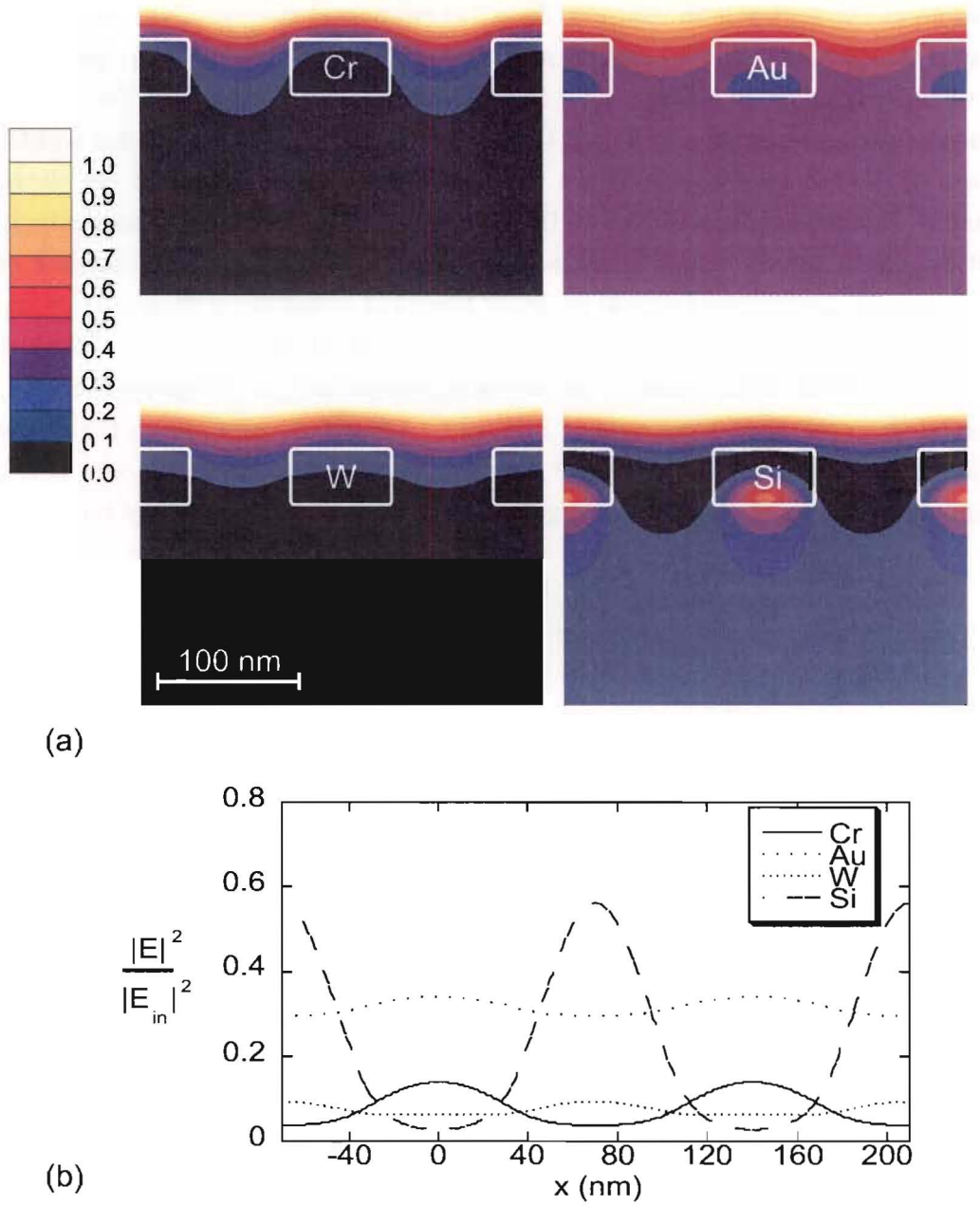


Figure 6.16 Normalised electric field intensity for a 140 nm-period grating illuminated with TE polarised light, for Cr, Au, W, and Si grating materials. (a) field maps with contour lines varying linearly from 0 (black) to 1.0 (white) in 10 linear steps. (b) Normalised intensity at the exit aperture of the gratings.

6.5.5 Effect of Grating Conductor Thickness

The behaviour of gratings has been found to be strongly dependent on the thickness of the grating conductors [126]. The effect of grating thickness on the contrast for an evanescent exposure was investigated for a 140 nm-period grating, illuminated with 436 nm, TM polarised light. The DOF at a contrast V = 0.5 ($DOF_{0.5}$) was used as a figure of merit to compare a range of Cr thicknesses from 10 nm to 70 nm. The results of these simulations are shown in Fig. 6.17 which plots the $DOF_{0.5}$ as a function of Cr thickness of the left-hand axis, and the maximum normalised intensity at the depth equating to $V = 0.5$ as a function of Cr thickness on the right-hand axis. Interestingly the $DOF_{0.5}$ is improved as the thickness of the metal is decreased. A number of things happen as the Cr thickness is reduced. Firstly the shadowing capability of the metal is reduced, particularly as the thickness reduces below the skin depth of chrome. However there is also an increase in the intensity, particularly in the aperture region, as illustrated on the right-hand axis of Fig. 6.17. Around the skin depth (δ) of the Cr, which is marked on the figure, the intensity reduces again. At thicknesses below the skin depth the Cr begins to have significant light penetration. Clearly the trend

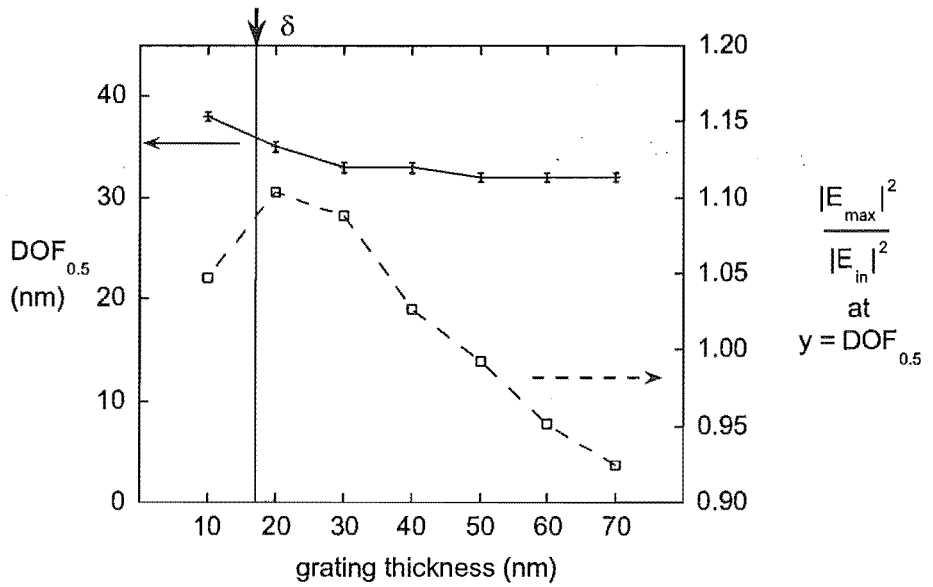


Figure 6.17 On the left axis the depth of field at $V = 0.5$ for a 140 nm period grating versus the grating thickness is plotted. On the right axis the maximum intensity at the depth equating to a $DOF_{0.5}$ is plotted versus the grating thickness.

of improving DOF with decreasing conductor thickness cannot continue indefinitely, but thinner grating conductors could not be simulated. When working with thinner chrome layers it becomes increasingly difficult to place adequate multipole expansions.

6.6 SIMULATION RESULTS FOR MULTI-LAYER ENFOL GRATINGS

Results for the metallic gratings suspended in air have provided valuable information about the possible resolution attainable by ENFOL-like exposures. To see how well these results might translate to the current experimental set-up, a more realistic multilayer model was constructed and simulated. The model is shown in Fig. 6.18. It

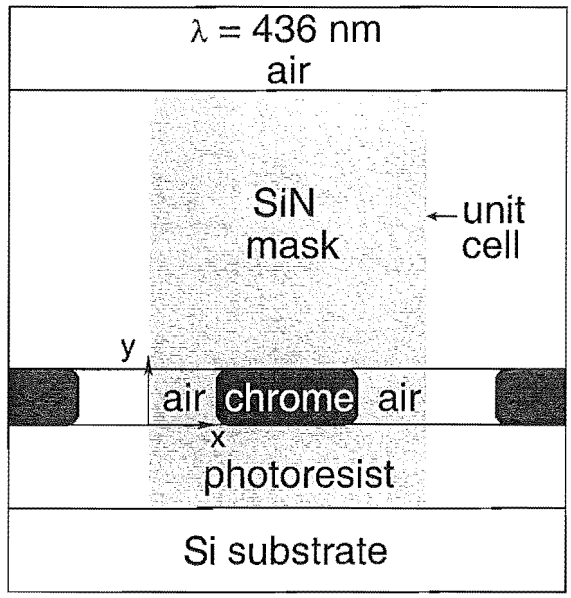


Figure 6.18 Model for multi-layer ENFOL.

incorporates a 200 nm thick SiN mask ($\varepsilon = 4$), a 40 nm thick Cr/air grating, a 60 nm thick photoresist layer ($\varepsilon = 2.56$), and an effectively infinitely thick silicon substrate ($\varepsilon = 22.6 + i1.6$). Material properties are taken from physical tables [94,131] at 436 nm.

The results for a 140 nm-period grating simulated with 436 nm incident light are illustrated in Fig. 6.19 for TM (Fig. 6.19(a)) and TE polarised light (Fig. 6.19(b)). They show similar characteristics to the models shown in Fig. 6.5 for Cr in air. Strong field enhancements exist at the grating corners for the TM polarisation, and again improved contrast is evident for TM in comparison to TE polarised light. Contrast in the TM exposure extends approximately 40 nm into the resist, and from the plot, one might expect an underexposed profile at the resist/Si interface.

A direct comparison of the 140 nm-period grating described above and that of a simple grating, again 140nm period, but this time suspended in photoresist is presented in Fig. 6.20. The first two line plots at $y = 0$ and $y = -20$ nm are very similar for the Cr in resist (Fig. 6.20(a)) and multilayer ENFOL model (Fig. 6.20(b)). The edge enhancements at the aperture are somewhat higher for the multilayer ENFOL

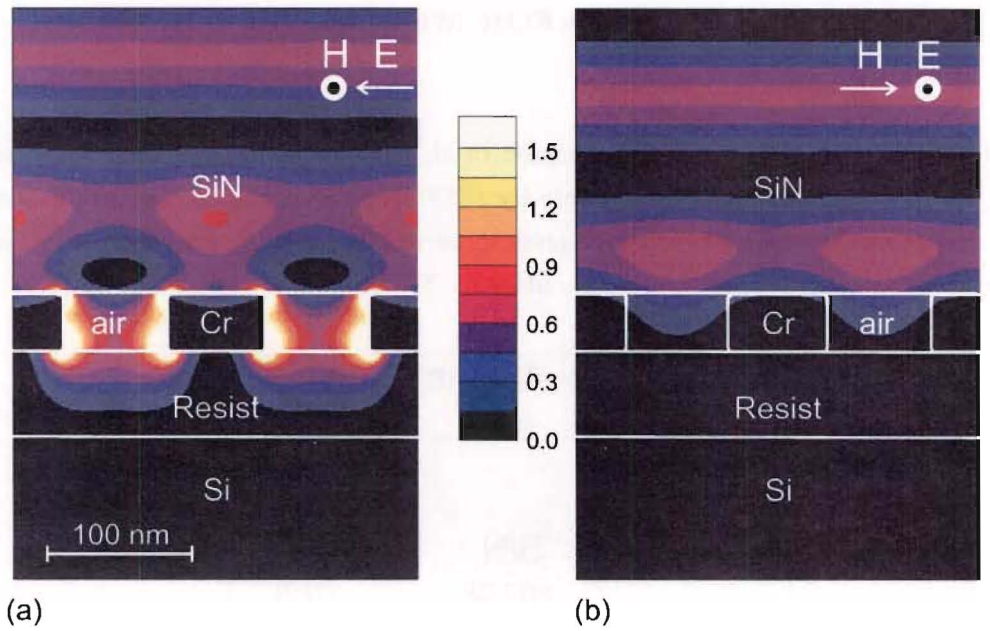


Figure 6.19 Normalised intensity distribution for a 140nm-period, 40nm-thick Cr grating on a 200nm thick SiN membrane. The exposure is made into 60 nm of photoresist ($n = 1.6$) on a Si substrate. The grating is illuminated from above by 436 nm linearly polarised light with (a) TM polarisation and (b) TE polarisation (electric field E and magnetic field H directions are indicated). Contour plots of the normalised electric field intensity are shown $|E|^2/|E_{in}|^2$, where E_{in} is the incident electric field. The scale varies linearly from 0 (black) to 1.5 (white) in 10 linear steps.

model, due to the Cr corners being in air as opposed to photoresist. In the photoresist medium the radiused edge appears less sharp with respect to the effective wavelength in photoresist. By $y = -20$ nm it is difficult to distinguish any difference between the plots.

The most important differences occur close to the resist/Si interface situated at y = -60 nm in Fig. 6.20(b). The electric field intensity is reduced here, to around 0.05 compared to around 0.2 for the suspended grating. This is due to the π phase change induced in the reflected wave as it encounters the silicon which has a higher refractive index than the photoresist. This results in destructive interference producing an intensity null. The reduction in intensity is undesirable from a lithography standpoint, as it means that the bottom of the photoresist will remain undeveloped. A more desirable configuration would have the substrate impedance-matched to the photoresist to prevent standing waves in the photoresist. These simulation results rationalise part of the experimental procedure for subtractive pattern transfer described in Section 5.3. A de-scum O_2 etch was found necessary to remove the residual resist layer present following exposure and development, prior to the Si etch.

Another interesting feature in the line-plots is the intensity inversion, where the intensity becomes greater in the region behind the opaque region of the grating than behind the aperture. This is visible at the y = -80 nm in Fig. 6.20(a) and y = -60 nm

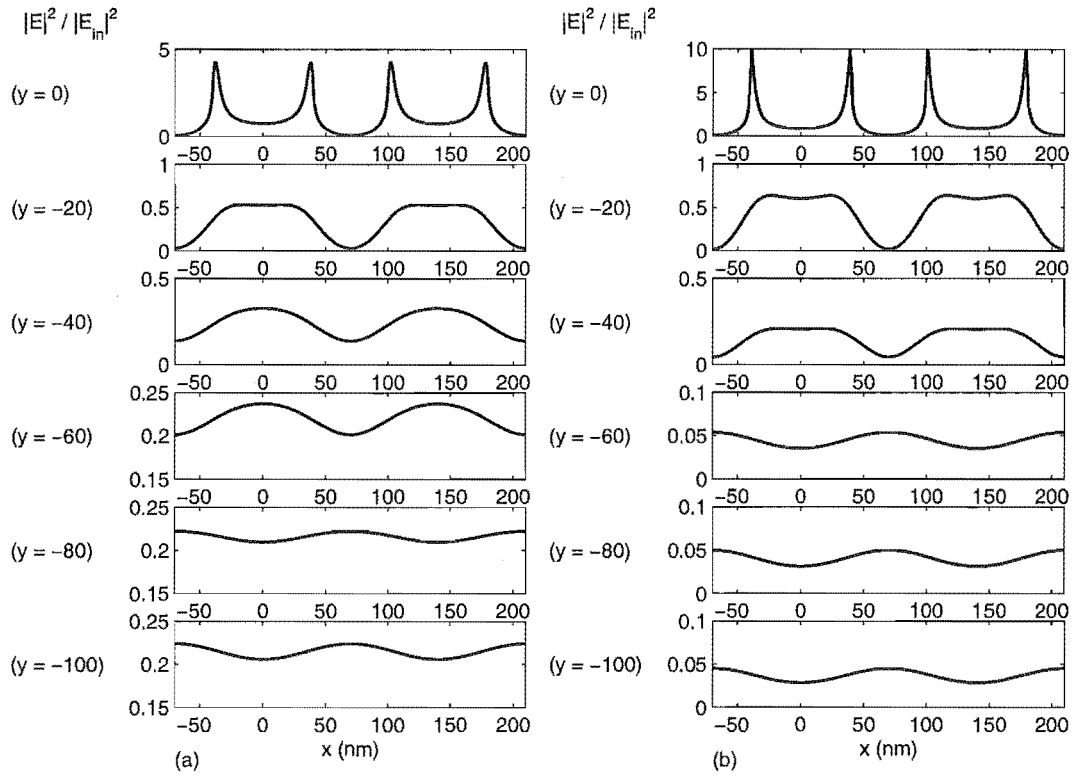


Figure 6.20 Normalised intensity profiles for 140 nm-period grating for TM polarised light for (a) a suspended grating in photoresist, (b) multilayer ENFOL model. Profiles are taken parallel to the grating direction. The 40 nm-thick grating spans from $y = 40$ nm down to $y = 0$ nm, hence the first cut ($y = 0$ nm) is at the exit plane of the grating then every 20 nm as indicated to the left of each intensity profile.

in Fig. 6.20(b). This phenomena is caused by constructive and destructive interference due to neighbouring apertures. This is an effect related to that first observed by Talbot [132] and later quantified by Rayleigh [133] known as the Talbot effect. At particular Talbot distances, a multiplication of the grating is observed.

The simulations of multilayer ENFOL models has been useful to highlight the issues of optically mismatched materials. They have also been able to confirm that the simplified models are adequate for understanding the main characteristics of the ENFOL exposure, as well as being more efficient to manipulate and simulate.

6.7 VALIDATION OF SIMULATIONS

Validation of the simulations was difficult as no rigorous analytic solutions exist that can be compared with the results to the authors knowledge. To some extent the errors in the boundary conditions output from MaX-1 serve as some measure of confidence in the results. Simulation results were not accepted using MMP until the maximum and average relative errors were in the range of 0.5% or less. If the errors were higher

refinements to the model were made. This means that at each matching point in the simulation, the boundary conditions agreed to within 0.5%, relative to the field magnitude at that point. Such low error is not sufficient in itself to confirm the validity of the results, although it does give a good indication along with direct observation of the resultant fields. Comparisons were also made of different MMP models with differing multipole placements and different ways of generating boundaries for the same geometry. These resulted in fields showing good agreement.

Further validation was obtained by comparing the MMP results to Gsolver results. GSolver is a commercially available electromagnetic simulator, that uses a different technique based on the rigorous coupled wave analysis (RCWA) described earlier in Section 6.2.3. GSolver became available near the end of this work. Simulations as close to the models simulated by MaX-1 as possible were constructed, although differences in the techniques meant the models were not identical. MaX-1 is able to represent circular boundaries, while GSolver is limited to straight-edged boundaries, with profiles being built up of separate lamellar layers producing a stepped profile. Two models were constructed in Gsolver and are illustrated in Fig. 6.21. The first (Fig. 6.21(a)) consists of a single lamellar layer, 40 nm thick comprising the Cr, and an abrupt transition to air. The second model, (Fig. 6.21(b)), shows an attempt to approximate the MMP model more closely by adding additional layers to the grating region. Nine layers are used, with the 30 nm thick centre layer surrounded above and below by four other layers, having the Cr air boundaries stepped to air at offset positions. In this way, a 5 nm radius is approximated by four 1.25 nm thick lamellar layers to smoothen the sharp corner.

Both the models illustrated in Fig. 6.21 were simulated with GSolver for a 140 nm-period grating, illuminated by 436 nm TM polarised light. Sixty diffracted orders were retained for the simulation, even though only the zeroth diffracted order is propagating. The high number of orders was necessary to obtain convergence for the TM polarisation to between two and three decimal places for the transmission and reflection coefficients. The GSolver output consists of the complex vector amplitudes of the transmission and reflection coefficients. Using the transmission coefficients (defined at the exit aperture of the grating), the fields below the grating were reconstructed. To do this a phasor projection was used, with the appropriate wave-vectors calculated using Equ. 2.10.

The results from the two GSolver models are illustrated graphically in Fig. 6.22, in which a comparsion with the MMP results is made. Overall, good agreement is demonstrated between all three models, with the agreement improving at greater distances below the grating. The greatest disparity in the results is evident at the exit aperture of the grating ($y = 0$ nm). The intensity peaks for the abrupt GSolver model have intensity peaks a factor of four times higher than MMP, while the layered GSolver model has peaks a factor of just under two times higher. These disparities can be explained in terms of the differences in the profile of the conductor corners. The model with

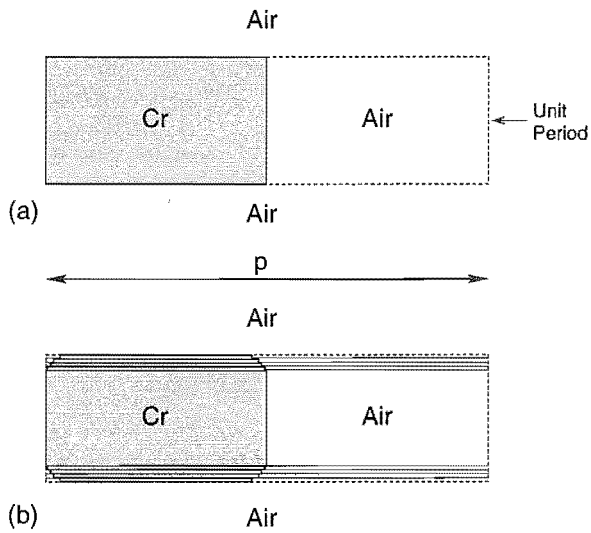


Figure 6.21 GSolver simulation models of Cr gratings suspended in air. (a) a three layer model, air substrate and superstrate with single layer Cr grating with an abrupt transition and (b) eleven layer model with air substrate and superstrate, and the Cr grating separated into nine layers, a central 30 nm thick grating with 1.25 nm thick layers above and below to round the corners. The Cr regions are presented shaded.

the sharpest corners, *i.e.*, the GSolver abrupt model, has the most intense fields, while the MMP model with the smoothest corners has the least intense fields. By increasing the number of layers for the GSolver model, it is anticipated that the differences would become even smaller in the near-field. Another marked difference at the grating exit aperture is the steepness of the sidewalls between the peak intensity and the Cr shadowed region. Both GSolver results show a more gradual skirted intensity profile that encroaches into the conductor further than MMP. This is likely to be due to a difference in the definition as $y \to 0$, in whether the material is metal or air. A fraction of a nanometre below zero, MMP too shows the same skirting. Note that the peaks in Fig. 6.22(a) are at slightly different positions on the x-axis. This is again due to differences in the conductor corner profiles. The peaks occur at a point consistent with the profiles greatest change in gradient. At 10 nm below the grating a higher intensity is noted for the GSolver abrupt model, however at depths below this, the profiles are indistinguishable.

These GSolver simulations confirm the validity of the MMP results. Experimental validation is of course most desirable, but more difficult to obtain quantitatively. Some qualitative confirmation of the simulation results with respect to the experimental work has been possible. The reducing depth of field with decreasing grating period has been clear when working with different resist thicknesses. Exposures into thick resist described in Section 5.2.3 has also emphasised this aspect, with the modulation in the resist decreasing with decreasing grating period. Variations in the pattern depths was also observed in Section 5.2.3 and is believed to be a function of grating duty cycle,

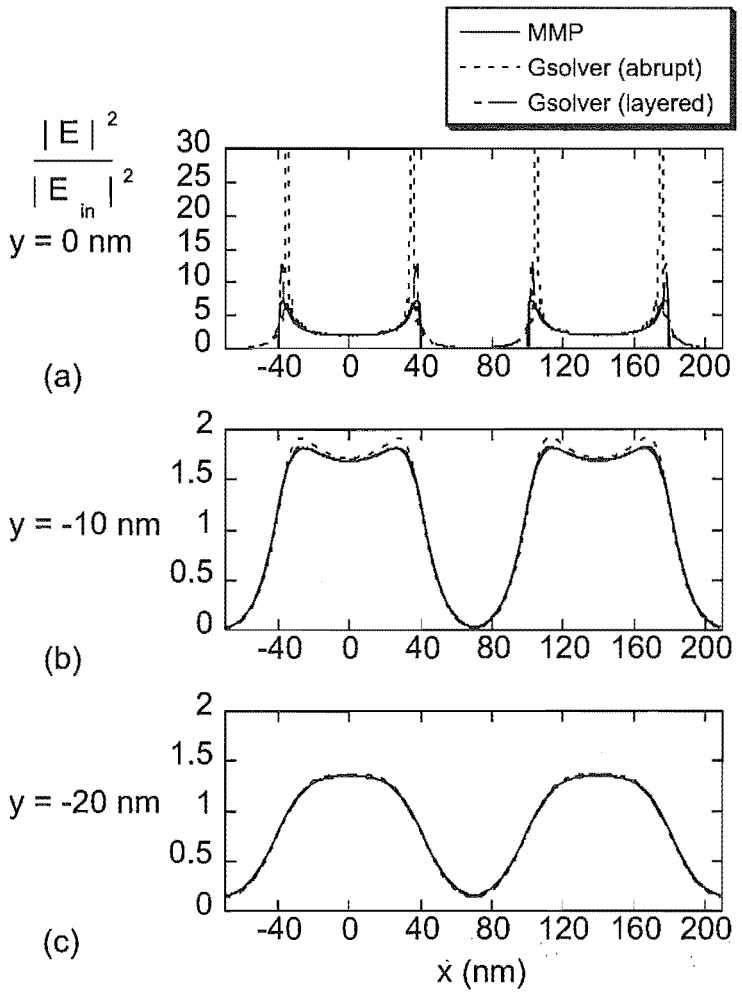


Figure 6.22 Comparison of near-field results between MMP and Gsolver models for a 140 nm-period grating illuminated by TM polarised light. Two Gsolver models are compared, a single-layer grating with an abrupt transition, and a model incorporating nine layers to approximate the radiusing used in MMP. The normalised intensity profiles are taken at (a) $y = 0$ (the exit plane of the grating), then at (b) 10 nm, and (c) 20 nm below the exit plane.

affecting the magnitude of T_0, the zeroth diffracted order, although this has not been proved rigorously.

6.8 EXTENSION OF RESULTS TO ARBITRARY FEATURES

Extension of the simulation results for sub-wavelength gratings to inform about exposures of arbitrary features is not possible. Gratings are a special and particularly useful example of sub-wavelength structure, where light of one linear polarisation is able to propagate due to one dimension (in our model the z) being sufficiently long compared to the wavelength. In this case there is ample intensity available for the exposure, although to obtain adequate contrast, operating in close contact with the grating is

essential. For geometries that do not contain a single dimension that is long compared to the wavelength, such as circular or rectangular apertures, no propagating polarisation exists. Analogies may be drawn with the gratings illuminated with TE light. For gratings illuminated with TE polarised light, there is very little intensity behind the grating for sub-subwavelength grating periods, and the range of the contrast is limited. This may make the exposure of sub-wavelength apertures difficult.

Not only is light transmission dependent on the geometry, it is also strongly related to the material properties. Recently Ebbesen and co-workers [65, 134] have demonstrated enhanced transmission through sub-wavelength holes in a high conductivity metal sheet. They report that the transmission, which exceeds that predicted by conventional theory, is due to surface plasmon resonances.

Investigation of the characteristics of exposures of sub-wavelength features with arbitrary shapes is vital for a more complete understanding of ENFOL-like exposures. The incorporation of different feature sizes and shapes to obtain a uniformly optimum exposure for all features is a challenging prospect. The scope of this task is huge, but anticipated to be rewarding.

6.9 SUMMARY

Simulations have been performed for ENFOL-like exposures of grating structures using the rigorous electromagnetic simulation technique MMP. Rigorous electromagnetic simulation techniques are the only way to investigate the fundamental limit of ENFOL and to support the ENFOL experimental work. Analytic solutions do not exist for the problems of interest, and the standard optical treatments for diffraction problems contain assumptions that are not valid for the sub-wavelength structures investigated in the near field.

Simulations of exposures of grating structures suggest there is no fundamental limit to the attainable resolution. Image contrast of $\lambda/20$ period gratings has been shown to exist, at distances of a few nanometres immediately below the grating. These resolutions are possible, despite the lateral dimensions of the grating absorbers being smaller than the skin depth of the Cr absorbers simulated. While the exposure depth is shallow, research into photoresist systems operating with only the top surface layer of atoms and molecules suggest that this may not be an unsurmountable issue.

The exposure of sub-wavelength gratings has been shown to be predominately due to TM polarised light, the linear polarisation preferentially transmitted by a wire-grid polariser. The intensity of the TE light has been found to have a negligible contribution to the exposure of sub-200 nm gratings at the 436 nm exposure wavelength. Enhanced evanescent components created from a TM exposure provide a high contrast image of the exposed sub-wavelength grating, for grating periods that have no propagating dif-

fracted orders. The evanescent nature of the diffracted components means the contrast has a limited range beneath the grating.

The depth of field, *i.e.*, the depth at which sufficient contrast is available for an exposure, is one of the critical parameters for an ENFOL-like evanescent exposure. Simulations of sub-wavelength gratings reveal a linear relationship between the depth of field and the grating period. From this an indication of a suitable resist thickness can be made for the grating period to be patterned. The decreasing depth of field with decreasing grating periods has also been confirmed experimentally. Simulation results predict sub-100 nm depths of field for grating periods below 300 nm, which is also consistent with experimental findings. For the exposure of masks with gratings of mixed periods, a resist thickness suitable for the smallest grating period will be required.

Near-field exposures appear to be strongly dependent on parameters such as duty cycle and the radius of curvature of the grating conductor corners. In particular, parameters that increase the zeroth order transmission coefficient of the TM polarisation appear to have a negative effect on the depth of field. For optimal exposure of mixed duty cycle patterns, this effect may need to be taken into account at the mask design stage. Again, operating at a resist thickness suitable for the pattern with the smallest depth of field will mitigate these effects, or alternatively selective application of attenuating mask materials.

High intensity edge enhancements have been noted at the conductor corners for the TM polarised exposures. They have a short range, generally limited to less than 10–20 nm. The edge enhancements have been found to be a function of the sharpness of the profile of the conductor corner. These are not anticipated to be problematic, although this may depend on the response of the particular resist system being exposed. The edge enhancements may be useful to obtain a period halving if suitable exposure conditions are used with a high resolution, ultra-thin resist.

Grating absorbers made from high conductivity materials with small skin depths and high absorbance appear to have the most desirable properties in terms of high contrast. Chrome appears to be a good material choice at the 436 nm exposure wavelength simulated. Tungsten has also been shown to be a potentially suitable alternative, with the processing advantage that it can be easily etched.

Useful insights into the defining characteristics of ENFOL exposures have been gained. ENFOL's potential as a high resolution lithography technique has been identified as well as potential pitfalls, namely the exposure variations that occur for mixed period structures, and for variations in other grating parameters such as duty cycle and the conductor edge profile. The examination of a variety of parameters that govern the performance of an ENFOL exposure should provide valuable feedback for optimising future experimental work.

Chapter 7

TRANSITION FROM EVANESCENT TO PROPAGATING EXPOSURES

7.1 INTRODUCTION

In Chapter 6 analysis of gratings was carried out for evanescent exposures, *i.e.*, exposures where all diffracted orders are evanescent. This regime has been of interest from an experimental standpoint, for evaluating one of the first reports of evanescent optical contact lithography [51–53] and to investigate the possible limit of the technique. This chapter investigates more generally the relationship between the effective illumination wavelength and grating period for evanescent to propagating exposures, concentrating in particular on the transition between these two regimes. At the transition, resonances result in a redistribution of energy between the different diffracted orders. These cause distinctive changes in the field intensity patterns which may not be desirable for lithography, where ideally a uniform exposure dose at the mask results in all features being exposed uniformly.

Interference phenomena that occur at the transition region could also be potentially useful [135]. A period halving of the grating is obtained which is appealing from an economic point of view as less tolerance for the mask manufacture is required. The transition region is investigated via simulation, and a new type of interference lithography is proposed that utilises enhanced evanescent diffracted orders. The effect of the grating material is investigated from the point of view of optimising the interference, but keeping in mind that it may be desirable to minimise this same interference for standard ENFOL.

In Section 7.2 of this chapter, large period gratings will be investigated and compared to the small period gratings studied in Chapter 6. Section 7.3 looks in detail at the transition region between large and small period gratings, where the effective wavelength is approximately equal to the grating period. Interference phenomena are discussed in Section 7.4, and an evanescent interference lithography (EIL) technique is proposed. The effect of different parameters such as grating material and absorbing resists upon the interference are discussed.

7.2 LARGE PERIOD GRATINGS

The traditional domain of contact lithography operates with illumination wavelengths smaller than the mask feature sizes to be replicated. In this domain light with propagating diffracted orders is responsible for the exposure. For the purposes of this work this equates to large period gratings which are defined here as $p \geq 2\lambda$. Contact lithography necessarily operates in the Fresnel near field. From scalar theory, this extends a distance of approximately

$$a^2/\lambda \tag{7.1}$$

from the diffracting aperture, where a is the aperture of the diffracting object [136]. For contact printing of gratings, a could equate to the size of the grating apertures on the mask. When the resist thickness is thinner than the depth of field of the smallest mask aperture, the field intensities provide an image that is a good representation of the mask, with some edge blurring due to diffraction. At depths greater than this, Fraunhofer diffraction effects result in the field intensities no longer appearing similar to the mask, and unsuitable for contact lithography. As can be seen from the relation defining the Fresnel/Fraunhofer boundary in Equ. 7.1, the Fresnel near field shrinks with decreasing mask feature size. The boundary however is not a rigid one, but rather a continuum with a gradual transition between the near and far fields.

Figure 7.1 shows an example of what the field intensities look like for a 1 μm-period Cr grating, suspended in air and illuminated by 436 nm light. The figure illustrates the field intensities for TM, TE and unpolarised illumination. The unpolarised intensity map is obtained by adding the intensities for the TM and TE illumination, and renormalising by halving the result. The boundaries of the 40 nm thick chrome conductors are outlined in white to make them distinct.

Noticeable differences are evident in the intensities beneath the grating for TM and TE illuminated gratings. This is because the ratio of $\lambda/p \approx 0.44$, is still relatively large, making the different boundary conditions for the two polarisations significant. Standing waves are visible above the gratings for both polarisations and field enhancements at the grating conductor corners are evident in the TM polarisation similar to what was observed for the small period gratings. Interestingly, the TE polarisation shows higher transmission intensity than for the TM case although both show good contrast in the first 500 nm below the grating. Below this, entering the far field regime, both the TE and TM polarisations show complex structure in the intensity, in contrast to the intensity plots for the sub-wavelength gratings simulated in Chapter 6.

Period halving can be observed, just over 1 μm below the grating in the TE polarisation. Spatial period division can be observed for periodic and quasi-periodic structures and is known as the Talbot effect, after its discoverer Fox Talbot [132]. At

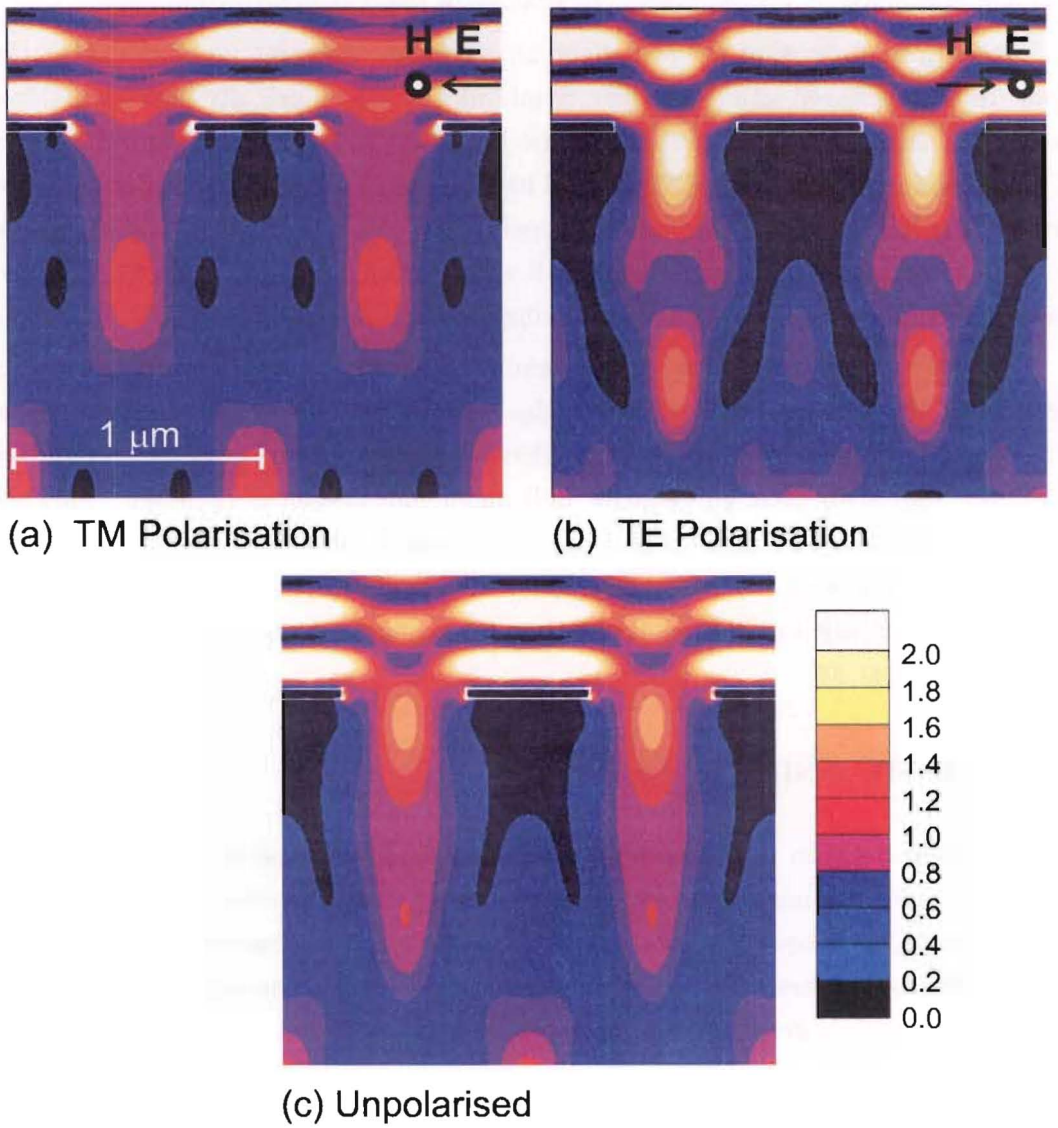


Figure 7.1 Normalised electric field intensity maps of a 1 μm-period Cr grating, illuminated by 436 nm light and (a) TM polarisation, (b) TE polarisation, (c) unpolarised. Intensity contours vary linearly from 0 (black) to 2.0 (white) in 10 linear steps.

particular distance planes (called Talbot distances) of $p^2/N\lambda$, the intensity pattern forms a self-replicating image with a period p/N, where N is the multiplicative factor. This phenomenon is due to constructive and destructive interference of the transmitted waves. In the TM polarisation however, Talbot effects are not as well defined and only a weak period halving can be seen near the bottom of the Fig. 7.1(a). A clear phase inversion is prominent 1.5 μm below the grating in the TM polarisation.

Spatial period division would appear an attractive prospect for microfabrication applications, where reductions in mask requirements is economically beneficial. Flanders *et al.* have demonstrated spatial period division with 196.8 nm period gratings halved to 98.4nm using x-rays [137]; period halving has also been observed at UV wavelengths [137]. Unfortunately, where spatial period division would be particularly useful, *i.e.*, where the mask period is getting small with respect to the wavelength, the effect is less well-defined, the sharpest images being obtained in the limit as $\lambda/p \to 0$ [138].

The unpolarised case in Fig. 7.1(c) demonstrates what contact lithographers believed to be the limit of contact lithography back in the 1970s. The grating's depth of field appears consistent with that predicted by the Fresnel/Fraunhofer boundary, $\approx$ 570 nm, indicating that gap control well under one micron is required. This was difficult to achieve over large areas with conventional chrome on glass rigid masks, however operating with conformable contact alleviates this problem. High contrast exposures are possible for features larger than the wavelength and this has been shown experimentally [43,53].

7.3 THE TRANSITION REGION, $\lambda \approx p$

The transition between an exposure with propagating waves and an exposure with evanescent waves, for normally incident light, occurs when the effective wavelength of illumination (λ/n) is equal to the grating period. This is equivalent to the cut-off of the $\pm$ 1 diffracted orders. The general equation for the cut-off wavelength of the mth diffracted order λ_m is given by

$$\left|\sin\theta_i + \frac{m\lambda_m}{pn}\right| = 1 \tag{7.2}$$

where n is the refractive index of the medium, θ_i is the angle the incident wave makes with the grating normal, and p is the grating period. Henceforth normal incidence is assumed, *i.e.*, $\theta_i = 0$.

To further investigate the change from propagating to evanescent exposures, the diffraction efficiencies for a grating were analysed over a range of wavelengths, centred around the cut-off of the $\pm$ 1 diffracted orders, λ_1, (note that for normally incident light $\lambda_1 = \lambda_{-1}$). The model simulated was a 270 nm-period Cr grating, suspended in a lossless photoresist with a real refractive index of 1.6. Illumination was at normal

incidence with the frequency varied from 6×10^{14} to 15×10^{14} Hz in steps of 1×10^{13} Hz; this equates to a wavelength range from 200–600nm. To take into account the frequency dependence of chrome's permittivity, polynomials were fitted to tabulated data [94]. To obtain the best fit for the real permittivity, the data was sectioned into two either side of the minimum of ε', with a second order polynomial fitted for the low frequencies and a third order polynomial fitted for the high frequencies. A third order polynomial was also fitted for ε'' with both plots showing close agreement to tabulated data, as shown in Fig. 7.2.

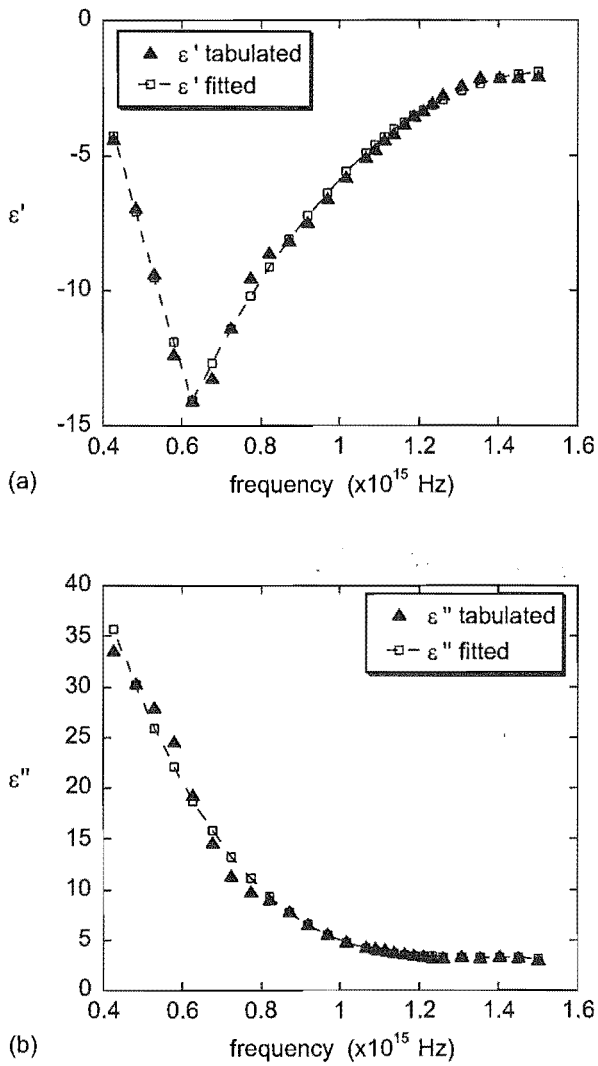


Figure 7.2 Comparison of tabulated and fitted data for the frequency dependent permittivity of chrome, (a) ε', (b) ε''.

The amplitude transmission coefficients t_m, and reflection coefficients r_m, give us important information about how the energy is distributed between the different diffracted orders. The transmission coefficient t_m is defined as the ratio of the amplitude

of the mth transmitted order and the incident electric field E_{tm}/E_{in}, and similarly r_m is defined as the ratio of the mth reflected order and the incident electric field E_{rm}/E_{in}. Traditionally the coefficients are far-field coefficients, so that once orders have reached cut-off, they have a value of zero. However, because the near field is of great interest in this work, a reference plane on the grating has been defined, to compare the relative strengths of the evanescent orders, whose value is now distance dependent. The grating exit aperture ($y = 0$ nm), has been chosen as the reference plane for the transmission coefficients, while the entry aperture ($y = 40$ nm), is the reference plane for the reflection coefficients. A schematic of what the transmission and reflection coefficients for a cut-off order might look like is illustrated in Fig. 7.3, with the reference planes defined.

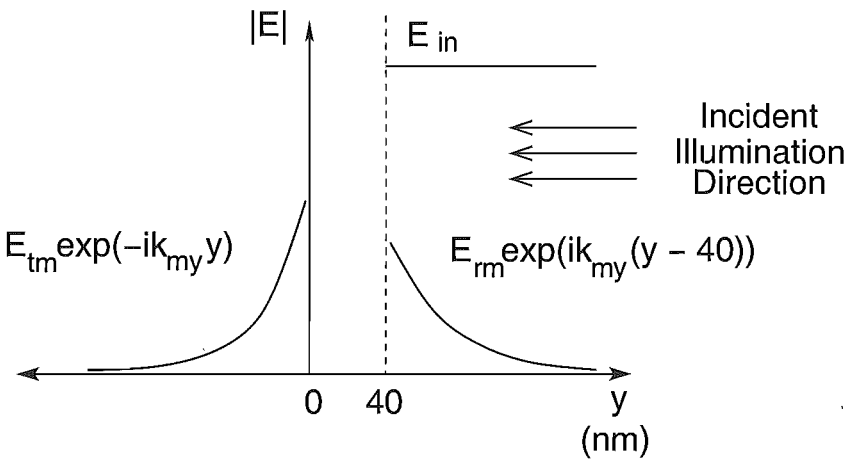


Figure 7.3 Schematic illustrating the distance dependence of the transmission and reflection amplitude coefficients of a cut-off diffracted order. The reference planes are defined for t_m at $y = 0$ nm and for r_m at $y = 40$ nm.

As the transmitted and reflected fields are computed at some distance from the grating it has been necessary to back-project these to the reference planes for the evanescent orders, taking into account the appropriate decay lengths of the different orders. The coefficients t_1 and r_1 are back-projected for wavelengths where $\lambda/p > 1$, and t_2 and r_2 for wavelengths where $\lambda/p > 0.5$, and assuming a decay constant of

$$k_{my} = i2\pi\sqrt{m^2/p^2 - n^2/\lambda^2}. \tag{7.3}$$

The amplitude transmission and reflection coefficients for $m = 0$, 1 and 2 are plotted as a function of the effective wavelength normalised to the grating period in Fig. 7.4 and Fig. 7.5 for the TM and TE polarisations respectively. Notable shifts in the distribution of energy can be seen in the transition region for both the TM and TE polarisations. For the TM polarisation near the cut-off of the first diffracted order (λ_1), the zeroth order transmitted wave t_0 shows a sharp dip in transmission. Concomitantly, the zeroth order reflected wave r_0, shows enhanced reflection. Both

the t_1 and r_1 coefficients show enhancements near λ_1, with r_1 peaking just before the peak in t_1. The second order coefficients t_2 and r_2 have significantly less energy than the lower orders and accordingly their effect on the field intensities is minor. The TE polarisation also exhibits strong resonances near λ_1, although these are less complex with a single peak seen for t_0, t_1, r_1, coincident with r_0 dipping sharply.

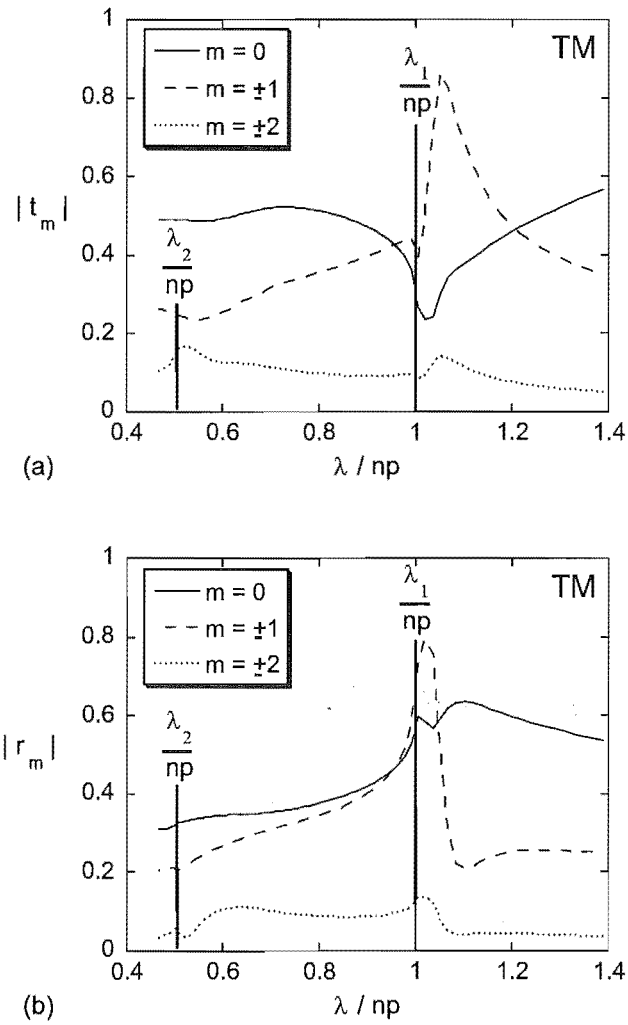


Figure 7.4 Zeroth, first, and second order diffraction (amplitude) coefficients plotted versus effective wavelength, normalised by the grating period for a 270 nm-period Cr grating. The grating is suspended in photoresist (n = 1.6) and illuminated with TM polarised light. (a) plots the transmission coefficients t_m and (b) the reflection coefficients r_m.

Figure 7.6 illustrates how the resonant diffraction phenomenon translates into field intensities for the grating at the first order cut-off, $\lambda_1 = 432$ nm, for both the TM and TE polarisations. For the TM incident polarisation, Fig. 7.6(a) shows a distinct halving of the grating period with high uniformity below the grating. The TE incident polarisation (Fig. 7.6(b)) gives a markedly different result. A checker-board structure results, with a periodicity equal to that of the grating.

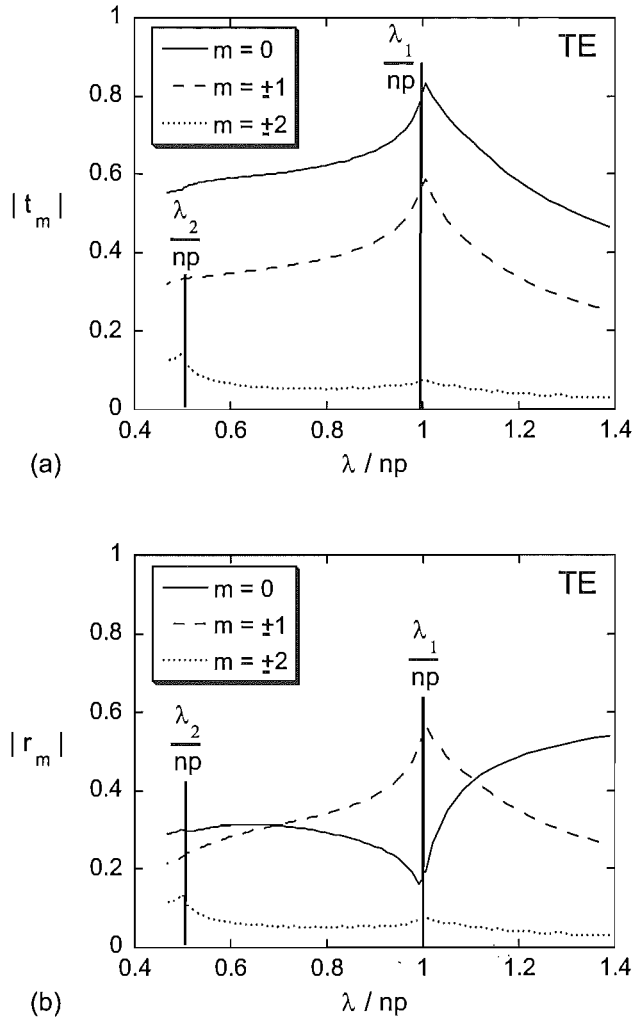


Figure 7.5 Zeroth, first, and second order diffraction (amplitude) coefficients plotted versus effective wavelength, normalised by the grating period for a 270 nm-period Cr grating. The grating is suspended in photoresist (n = 1.6) and illuminated with TE polarised light. (a) plots the transmission coefficients t_m and (b) the reflection coefficients r_m.

The form of the resultant intensities at cut-off *i.e.,* the period halving for the TM illumination and checker-board appearance for the TE illumination, can be explained using a simplified analysis considering only the zeroth and first order diffracted terms. At cut-off, from Equ. 2.10, Equ. 2.11 and Equ. 2.12 we find that for the first diffracted orders $k = k_{1x}$ and $k_{1y} = 0$. This indicates propagation of the first-order diffracted components in the x direction, the direction of the grating vector.

Firstly, an analysis for the TM model will be considered, neglecting all diffracted orders > 1. The zeroth transmitted component consists of an E_x component, propagating in the y direction. Note, as the zeroth diffracted order has $k_{0x} = 0$, $k_{0y} = k$. The ± 1 diffracted components have E_y components and are propagating in the x direction. The fields of these components are then in quadrature, and the components of the fields

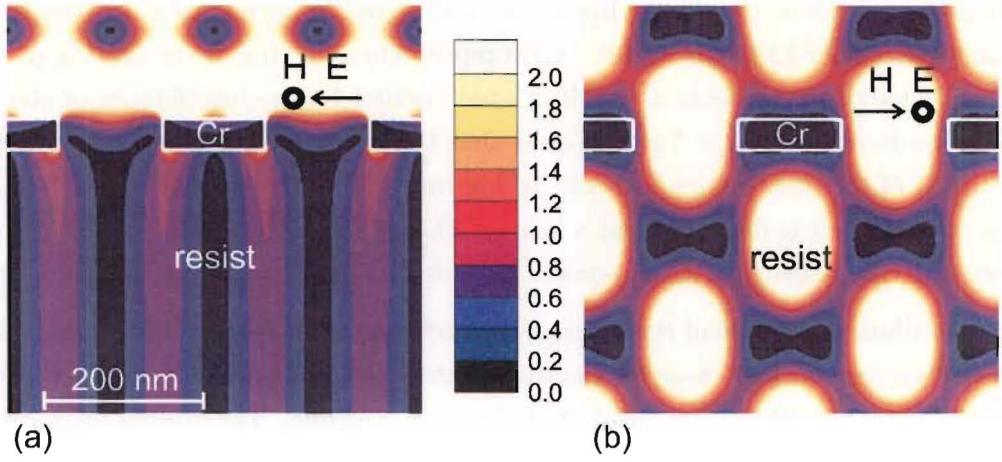


Figure 7.6 Normalised electric field intensity maps of a 270 nm-period Cr grating, suspended in photoresist (n = 1.6) and illuminated by 432 nm light for (a) TM polarisation, (b) TE polarisation. Contours vary linearly from 0 (black) to 2.0 (white) in 10 linear steps.

can be determined independently. The normalised electric field E_y can be expressed as

$$\frac{E_y}{|E_{in}|} = t_1 \cos(k_{1x}x - \omega t) + t_{-1} \cos(k_{-1x}x + \omega t), \tag{7.4}$$

two waves propagating in opposite directions. The normalised electric field E_x below the grating can be expressed as

$$\frac{E_x}{|E_{in}|} = t_0 \cos(k_{0y}y - \omega t). \tag{7.5}$$

The intensity then has the form

$$\frac{|E|^2}{|E_{in}|^2} = \left(\frac{E_y}{E_{in}}\right)^2 + \left(\frac{E_x}{E_{in}}\right)^2. \tag{7.6}$$

Assuming $t_1 = t_{-1}$ (consistent with normal incident light), and $k_{1x} = k_{-1x} = k = k_{0y}$ this simplifies to

$$\begin{aligned}\frac{|E|^2}{E_{in}^2} &= t_1^2 + \frac{t_1^2}{2}\cos 2(kx - \omega t) + \frac{t_1^2}{2}\cos 2(kx + \omega t) + \\ &\quad t_1^2 \cos 2kx + t_1^2 \cos 2\omega t + \\ &\quad \frac{t_0^2}{2} + \frac{t_0^2}{2}\cos 2(ky - \omega t).\end{aligned} \tag{7.7}$$

Integrating over one cycle, the time averaged form is,

$$\frac{1}{T}\int_{-T/2}^{T/2} \frac{|E|^2}{E_{in}^2} dt = \frac{t_0^2}{2} + t_1^2 + t_1^2 \cos 2kx. \tag{7.8}$$

This equation contains two unvarying terms, and a $\cos 2kx$ term. As k is equal to the grating wavenumber $2\pi/p$, the $\cos 2kx$ term represents a halving of the grating period. The period halving represents a standing wave created by the interference of the ± 1 diffracted orders. From Equ. 7.8 it is clear that the relative values of t_0 and t_1 define the contrast of the interference pattern. To improve the contrast of the period halving, a large t_1 coefficient is desirable and a low t_0. This can be seen in the limit as $t_0 \to 0$; the intensity has a raised cosine response $t_1^2(1+\cos 2kx)$ that has perfect intensity nulls.

At the illuminated side of the grating, distinct period halving is also evident. This is due to the same mechanism as for the transmitted case although with higher contrast due to the larger amplitude of r_1 compared to t_1 at $\lambda = 432$nm. The interference pattern becomes dots not lines because of the standing wave resulting from interference of the incident and reflected waves.

Now to analyse the TE illuminated case, again at cut-off, the ± 1 diffracted orders propagate in the x direction, but with an E_z component. The zeroth transmitted order propagates in the y direction, also with an E_z component. As the fields have the same direction, a scalar addition can be performed,

$$\frac{E_z}{E_{in}} = t_0 \cos(k_{0y}y + \omega t) + t_1 \cos(k_{1x}x - \omega t) + t_{-1} \cos(k_{-1x}x + \omega t). \tag{7.9}$$

Substituting $t_1 = t_{-1}$ (consistent with normal incident light), and $k_{1x} = k_{-1x} = k = k_{0y}$ the intensity has the form

$$\begin{aligned} \frac{|E_z|^2}{E_{in}^2} &= \frac{t_0^2}{2} + \frac{t_0^2}{2}\cos 2(ky + \omega t) + t_0 t_1 \cos(k\dot{y} + kx + 2\omega t) + \\ &\quad t_0 t_1 \cos(ky - kx + 2\omega t) + t_0 t_1 \cos(ky + kx) + t_0 t_1 \cos(ky - kx) + \\ &\quad \frac{t_1^2}{2} + \frac{t_1^2}{2}\cos 2(kx + \omega t) + \frac{t_1^2}{2}\cos 2(kx - \omega t) + \frac{t_1^2}{2}\cos 2kx + \frac{t_1^2}{2}\cos 2\omega t. \end{aligned} \tag{7.10}$$

If the time average is taken over one cycle,

$$\begin{aligned} \frac{1}{T}\int_{-T/2}^{T/2} \frac{|E_z|^2}{E_{in}^2} dt &= \frac{t_0^2}{2} + t_1^2 + t_0 t_1 \cos(ky + kx) + t_0 t_1 \cos(ky - kx) + t_1^2 \cos 2kx \\ &= \frac{t_0^2}{2} + t_1^2 + 2t_0 t_1 \cos ky \cos kx + t_1^2 \cos 2kx. \end{aligned} \tag{7.11}$$

From the time averaged result, an x-y dependency is evident, with a periodicity in both x and y, equal to that of the grating period. This is consistent with the checker-board appearance of the intensity map. A spatial frequency doubled term is also present, but from the transmission coefficients in Fig. 7.5(a), it is clear that it is not dominant.

High transmission of the zeroth and ± 1 components for the TE polarisation provide a high contrast pattern in comparison to the TM polarisation. One can imagine that

if a material was chosen with a refractive index strongly dependent on exposure, a 2-D equivalent of a photonic band-gap (PBG) structure [139, 140] could be produced in the x-y plane. For a true 3-D PBG, an additional exposure would be required. Some compromise may be required to ensure the refractive index change is not so large that it shifts the exposure out of resonance, by changing the effective wavelength dramatically. Photoresist is one example of a material whose refractive index changes with exposure (bleaching), however the magnitude of the change is often small.

7.4 EVANESCENT INTERFERENCE LITHOGRAPHY

The period halving shown in previous simulations for gratings with $p = \lambda/n$ for TM incident light could yield potentially useful applications. From the transmission coefficients plotted in Fig. 7.4(a), a grating resonance showing enhanced ± 1 diffracted orders with respect to the zeroth order suggests an attractive possibility of a period halving with improved contrast. As the resonance occurs at wavelengths for which the ± 1 diffracted orders are cut-off, the period halving is due to interference of evanescent diffracted orders. Use of this effect for lithography is proposed as a new technique named evanescent interference lithography (EIL) [135]. This section discusses the principle of interference lithography, and reports on some interference schemes utilising propagating waves for microfabrication. The potential of EIL is then investigated via simulations.

Interference lithography, also known as holographic lithography, refers to the lithographic exposure resulting from two or more interfering waves. In its simplest form a grating can be produced by two coherent interfering plane waves intersecting at an angle of 2θ to produce a grating of period $p = \lambda/(2sin\theta)$, where λ is the illumination wavelength. If grazing incidence is assumed, ($\theta = 90°$), a minimum p of $\lambda/2$ can be obtained, which is also considered the resolution limit for photolithography with propagating light.

Interference lithography is an attractive candidate from a number of standpoints; it is a simple means of generating periodic structures, it has good depth of field, and is able to pattern large areas with high resolution. The main disadvantage is its limited flexibility, as it is only able to pattern periodic or quasi-periodic structures. Having said that, there are a large number of applications that incorporate periodic structures, including filters, distributed feedback (DFB) and distributed Bragg reflector (DBR) lasers, narrow-band filters for wavelength division multiplexing, and anti-reflection structures. For some of these applications, precise control over the period and duty cycle are essential.

A number of variants of interference lithography have been reported for microlithography applications. Generally only a laser (or narrow-band source) and a phase mask is required. The phase mask diffracts the light into different orders which then interfere.

In one technique, monochromatic light illuminates a phase grating at an angle θ_i from normal. For a specific wavelength λ, a combination of the angle of the incident light and the period of the phase grating are chosen so that only the 0th and -1st orders propagate. The mth order is a propagating order if $|k_{mx}| < 2\pi n/\lambda$ is satisfied, where k_{mx} is the wavenumber of the mth order in the x direction, $k_{mx} = 2\pi\lambda \sin\theta_i + m2\pi/p$. The phase grating depth and angle are designed to give close to equal intensity in both orders. Tennant *et al.* have used this technique in the near-field for the generation of gratings that have an aperiodicity incorporated, for the fabrication of a $\lambda/4$-shifted DFB laser with a single exposure [141,142].

Another technique that also uses a phase grating is able to produce an interference pattern with half the period of the phase grating. This is achieved using monochromatic light, normally incident on a phase grating, with the grating depth selected to suppress the zeroth diffracted order. The period of the grating is such that only the ± 1 diffracted orders are propagating. These two orders interfere, resulting in a period halving. This technique is used to manufacture one-dimensional Bragg gratings in the cores of optical fibres [143]. Suleski *et al.* have shown that the same technique can be used for creating more complex two-dimensional gratings and radial gratings [144]. Both the above techniques have demonstrated near-field exposures using a conventional mercury arc-lamp with a filter to generate quasi-monochromatic light [144,145] as well as using laser sources. Multiple-exposure interferometric lithography has been investigated by Brueck *et al.*, which increases the dimensionality of an exposure to allow more complex two-dimensional structures to be fabricated [146,147]. A number of sequential interferometric exposures using a coherent laser are made, with either one or many variables changed between each exposure. There are a number of degrees of freedom, including the insertion of a different period phase or amplitude mask, a rotation of the sample, or a change in phase by shifting the sample relative to a phase/amplitude mask, or a change in wavelength or angle of incidence.

For all of the aforementioned interference lithography techniques, propagating waves are interfered. Similar effects can also be obtained using evanescent waves, with increased intensities, if frequency dependent grating resonances are exploited. Revisiting the transmission coefficients for the TM polarisation in Fig. 7.4, a distinct peak in the t_1 coefficient is notable at approximately 454 nm, which is past this order's cut-off and is therefore in the evanescent regime. By selecting an operating point based on the transmission coefficients, one might expect to be able to optimise the period halved exposure, which from Equ. 7.8 appears to be when t_1 is a maximum and t_0 a minimum. The intensity field maps shown in Fig. 7.7 confirm a resonance at 454 nm compared to 432 nm, however, as the t_1 component is not propagating, there is a limited depth of field. Whereas at cut-off ($\lambda = 432$ nm), the t_1 component appears to have an infinite depth of field, at 454 nm very little contrast is evident after 200 nm due to the ± 1 orders being evanescent.

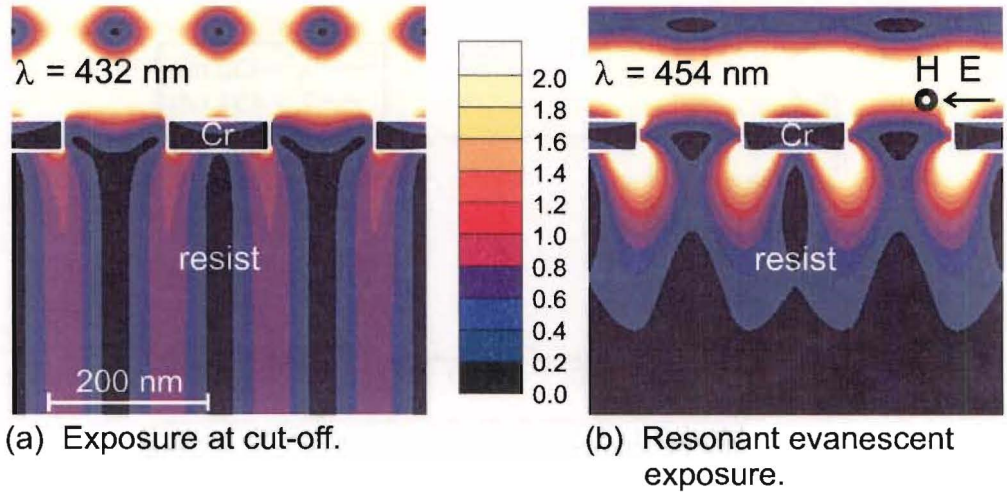


Figure 7.7 Normalised electric field intensity maps of a 270 nm period Cr grating, suspended in photoresist (n = 1.6) and illuminated by TM polarised light (a) $\lambda = 432$ nm (cut-off), (b) $\lambda = 454$ nm. Intensity contours vary linearly from 0 (black) to 2.0 (white) in 10 linear steps.

The enhancement of the interference pattern at 454 nm can be explained in terms of surface plasmon enhancement; such enhancements occur when there is momentum matching between the incident light photons and oscillating charge on the grating [65]. Figure 7.8 demonstrates more quantifiably both the intensity enhancement and the effect of the reduced depth of field for EIL. Figure 7.8 shows line-plots that compare the intensities for an exposure at cut-off ($\lambda = 432$ nm), and at resonance ($\lambda = 454$ nm), at different depths beneath the grating. At the exit aperture, and 10 nm below the grating, illumination with 454 nm light results in peaks enhanced by a factor greater than four compared to illumination with 432nm light. Better contrast for EIL is evident in the first 50–100 nm compared to the case at cut-off. This enhancement continues to be greater at 454 nm until 150 nm below the grating, however the intensity nulls in the apertures of the grating (x = 0 nm, x = -270 nm) are not as deep compared to the 432 nm illumination, reducing the interference contrast and making the 432 nm exposure more desirable in terms of pattern uniformity. However, if high intensity in a thin photoresist were required, the EIL regime is better.

Figure 7.9(a) shows the wavelength dependence of the peak intensity in the interference pattern at various depths below the grating. Note that the exit aperture results are omitted as they swamp the other data. Below cut-off (λ_1), the peak intensity is relatively insensitive to the wavelength. The picture changes dramatically above cut-off, with a strong enhancement in the peak intensity between 450–460 nm. At 10 nm below the grating the maximum peak intensity is almost five times greater than the incident intensity. Deeper than 100 nm, the peak intensity drops beneath that of the propagating wavelengths and there is no longer any advantage to operation in the evanescent regime. The enhancement decays strongly with distance, and the wavelength that gives

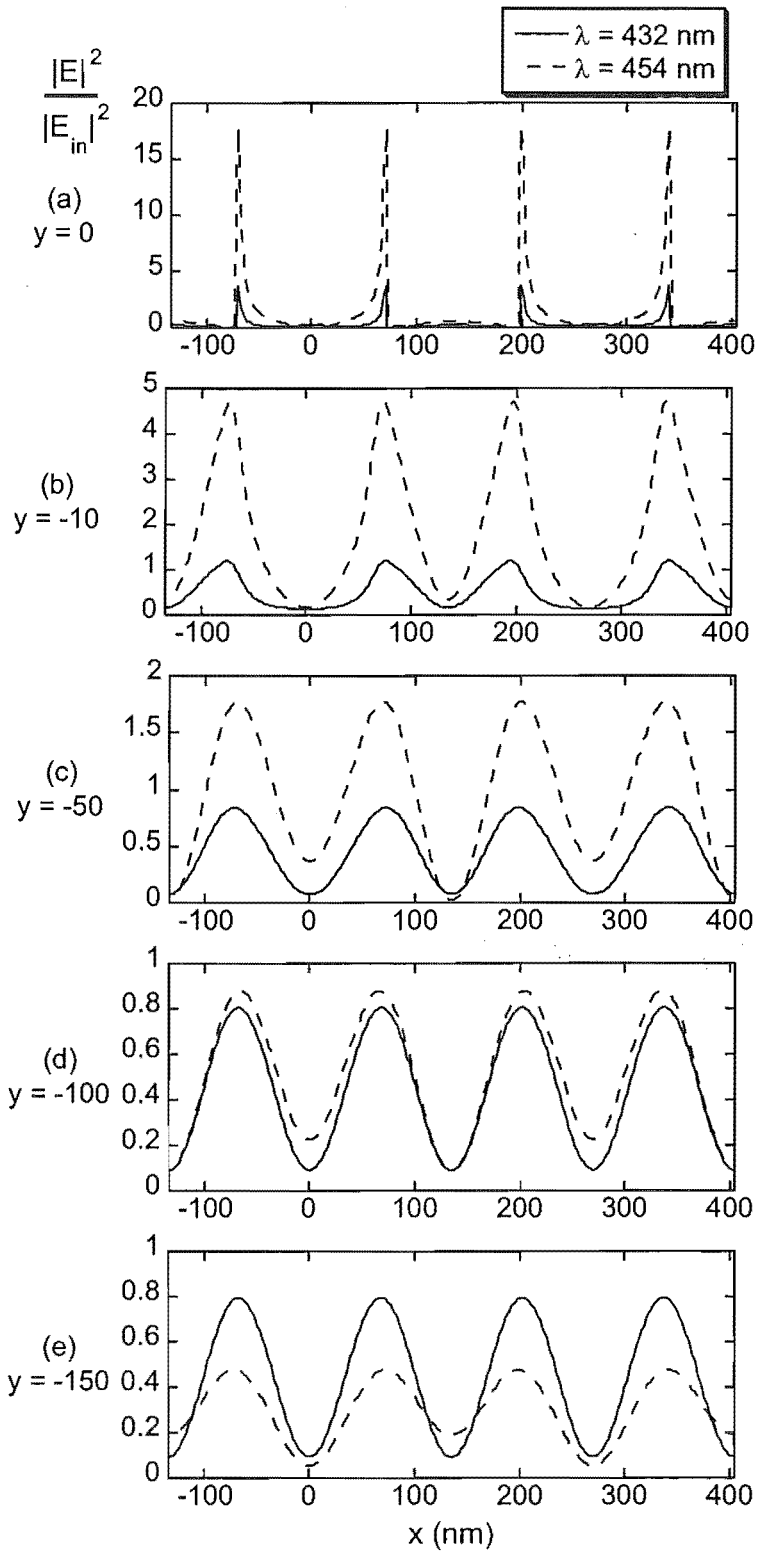


Figure 7.8 Line-plots of the normalised electric field intensity for a 270 nm-period Cr grating, suspended in photoresist (n = 1.6) and illuminated by 432 nm light TM polarised light at (a) $y = 0$ nm (the exit aperture), (b) $y = -10$ nm, (c) $y = -50$ nm, (d) $y = -100$ nm, (e) $y = -150$ nm.

the largest peak intensity becomes progressively shorter with increasing depth below the grating. This is a result of the increasing decay constant k_{1y} with wavelength where $k_{1y} = i2\pi\sqrt{1/p^2 - n^2/\lambda^2}$.

Just as important as the maximum intensity is the quality of the intensity null introduced below the grating apertures due to the interference. Figure 7.9(b) plots the intensity at the centre of the aperture versus wavelength at different depths below the grating. Prior to cut-off there is only a weak null, due to the significant t_0 component. This is followed by a region of wavelengths $\approx$ 430–450 nm with a deep null, before the null starts to disappear again at longer wavelengths. The fading out of the interference null with increasing wavelength, is a consequence not only of moving out of the resonant zone and t_0 becoming more significant, but also as a result of the reduced range over which the evanescent diffracted orders can interfere. This is why period halving is not evident for $\lambda \gg p$.

An image contrast factor V can be determined which describes the visibility or contrast of the interference. Actually, there are two contrasts, one with respect to the null in the apertures of the grating, and another with respect to the null under the centre of the grating absorbers. For EIL the V is defined as

$$V = \frac{I_{max} - I_a}{I_{max} + I_a}, \tag{7.12}$$

where I_{max} is the maximum intensity below the conductor edges, and I_a the intensity in the apertures. The contrast with respect to the other null below the grating conductors is omitted because at the depths of interest it is less sensitive to wavelength shifts. Figure 7.9(c) plots the contrast as a function of wavelength at different depths beneath the grating. This reinforces that improved contrast is obtained at wavelengths in the evanescently enhanced region for the first 50 nm beneath the grating. At cut-off (λ_1), the contrast at all depths converges to a single V value, consistent with the infinite depth of field case ($k_{1y} = 0$) evident in Fig. 7.7(a).

Another important point to note is that for successful EIL, good polarisation control is essential. A high extinction ratio is required to reduce the TE polarisation to protect the intensity null in the grating apertures, this is especially important as the TE polarisation is resonantly enhanced at cut-off. Operating deeper in the evanescent regime works in EIL's favour with respect to the polarisation, as the TE resonance starts to fade for Cr gratings. Threshold phenomena associated with the second order cut-off λ_2 are less prominent than at λ_1, suggesting a similar interference effect to obtain a period quartering is not feasible.

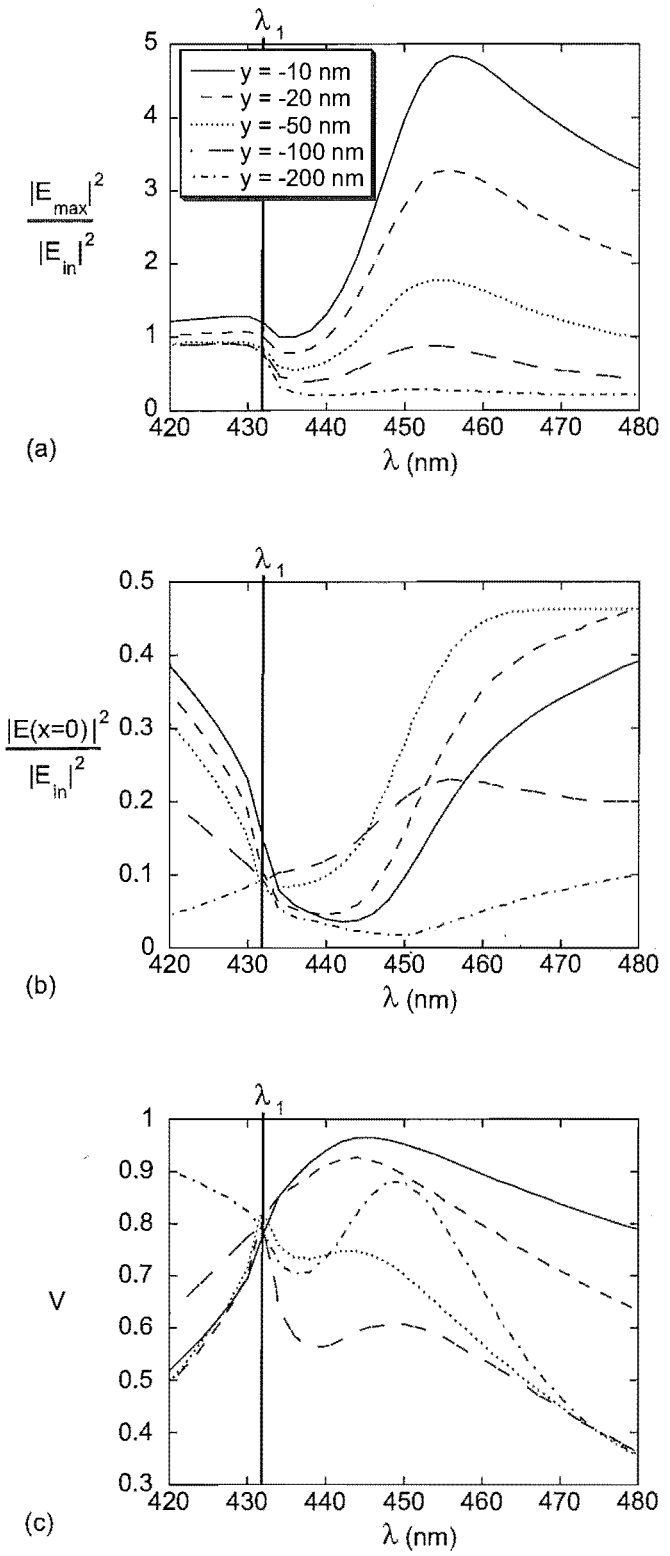


Figure 7.9 Field intensity of the interference as a function of illumination wavelength (TM), (a) peak intensity, (b) intensity at the centre of the aperture, at the exit plane and various distances below. In (c) the contrast of the additional interference null at the grating apertures as a function of illumination wavelength (TM), at the exit plane and various distances below. The first order cut-off is indicated.

7.4.1 Effect of Conductor Material

Period halving is expected around the cut-off of the ± 1 orders for gratings of other metallic materials illuminated by TM light owing to the interference of the surface waves created. The resonantly enhanced behaviour, however, is expected to be strongly governed by the material properties of the grating and the surrounding medium. The effect of the absorber material on EIL was investigated by comparing gold (Au), well known as a good conductor, and silicon (Si), a strongly polarisable dielectric, with Cr gratings in the resonance regime. Similar resonant effects were looked for to see if improved enhancements could be obtained with other materials. It has been reported that threshold phenomena connected with energy redistribution are often observed at the cut-off of different orders [148], however, this redistribution is strongly dependent on the material properties and how the boundary conditions are satisfied for the different polarisations.

Simulations for Au and Si gratings were performed using GSolver [114] to investigate the diffraction efficiencies as a function of wavelength. GSolver was used due to the ease of creating user defined wavelength dependent models for the refractive index. Polynomials were fitted to tabulated data [94] to represent the wavelength dependence of the complex refractive indices of Au and Si. For Au, a ninth order polynomial was fitted for n', and a seventh order polynomial for n''; the fitted data is compared to tabulated data in Fig. 7.10. For Si, both n' and n'' were fitted with tenth order polynomials as shown in Fig. 7.11. MaX-1 can theoretically do this too, but to obtain an accurate match to the tabulated values, the order of the polynomial required to fit the permittivity exceeded the text input capabilities of the package. A nine layer grating model, similar to that used in the Section 6.7, was used with a 270 nm period and surrounded by a lossless photoresist ($n = 1.6$). A sweep was then made of wavelengths from 200 nm to 700 nm in 5 nm steps to determine the diffraction efficiencies, with 30 orders maintained for the calculation.

Figure 7.12 plots the zeroth, first and second order transmission and reflection coefficients for an Au grating illuminated with both linear polarisations. For the TM polarisation, no distinct energy re-distribution is noted at λ_1 in either the transmission or reflection coefficients. For wavelengths longer than around $1.2\lambda_1$ a broad resonance is evident with a dip in the t_0 and an increase in t_1, with these having the largest disparity at around $1.5\lambda_1$ or $\lambda = 645$ nm. Comparing these plots to those of the Cr grating in Fig. 7.4, there are some similar characteristics, namely the t_0 dip, and the t_1 peak, however for the Au grating the resonance appears shifted further into the evanescent region and broadened. For EIL, the Cr appears to have more favourable properties as the closer the resonance is to cut-off, the better the depth of field.

The coefficients for the TE polarised illumination plotted in Fig. 7.12(c) and Fig. 7.12(d) show a much closer resemblance in shape and magnitude to those of the Cr

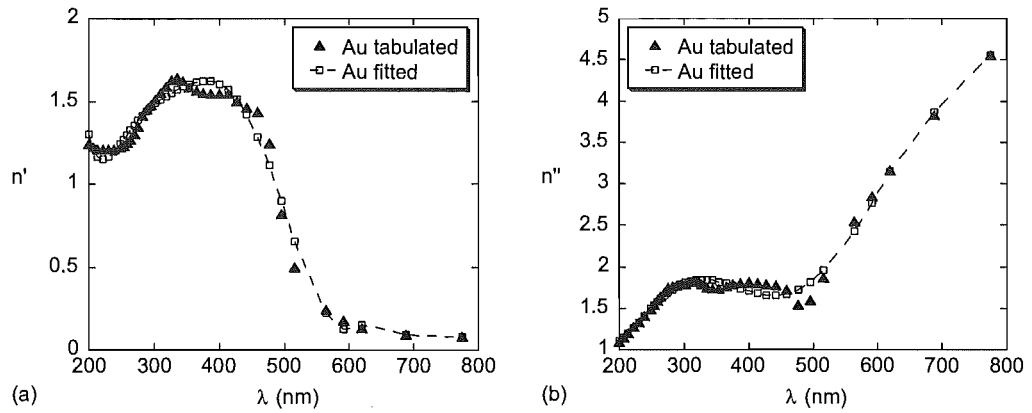


Figure 7.10 Comparison of tabulated and fitted data for the wavelength dependent refractive index of gold, (a) represents the real refractive index (n') which was fitted with a 9 order polynomial, while (b) represents the extinction coefficient (n''), fitted with a 7 order polynomial.

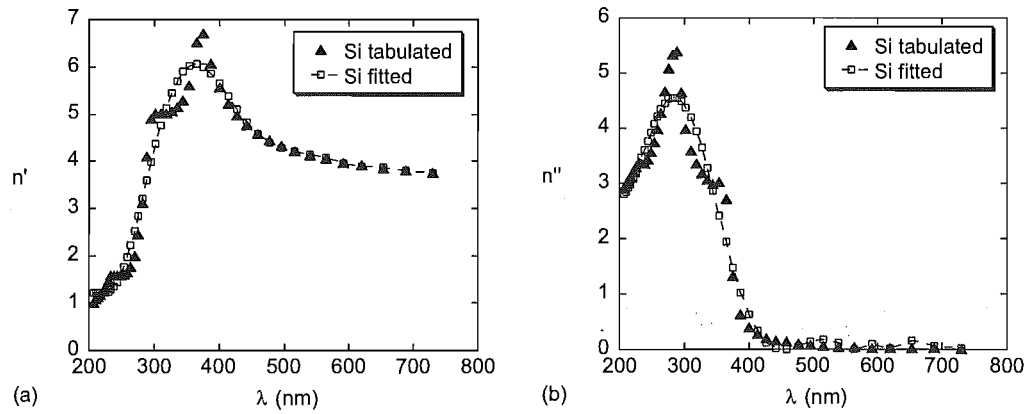


Figure 7.11 Comparison of tabulated and fitted data for the wavelength dependent refractive index of silicon, (a) represents the real refractive index (n'), while (b) represents the extinction coefficient (n''), both fitted with 10 order polynomials.

grating (Fig. 7.5), suggesting that the TE polarisation is less sensitive to the properties of the grating material.

Figure 7.13 illustrates the 2-D intensity maps for the Au grating from Fig. 7.12 at two of the wavelengths of interest (a) at cut-off, and (b) at 645 nm, the wavelength corresponding to the minimum of t_0. As expected, the period halving is much weaker at cut-off as the t_0 component is a factor of three greater than t_1. At the resonance in Fig. 7.13(b) there is no period halving to speak of, just edge-enhancements at the grating corners, as the range of the interference has been severely reduced. This last figure resembles more the results for gratings with periods much less than the wavelength. Operating at resonance here has some additional disadvantages, the enhancement spreads a long way laterally under the Au, reducing the shadow region.

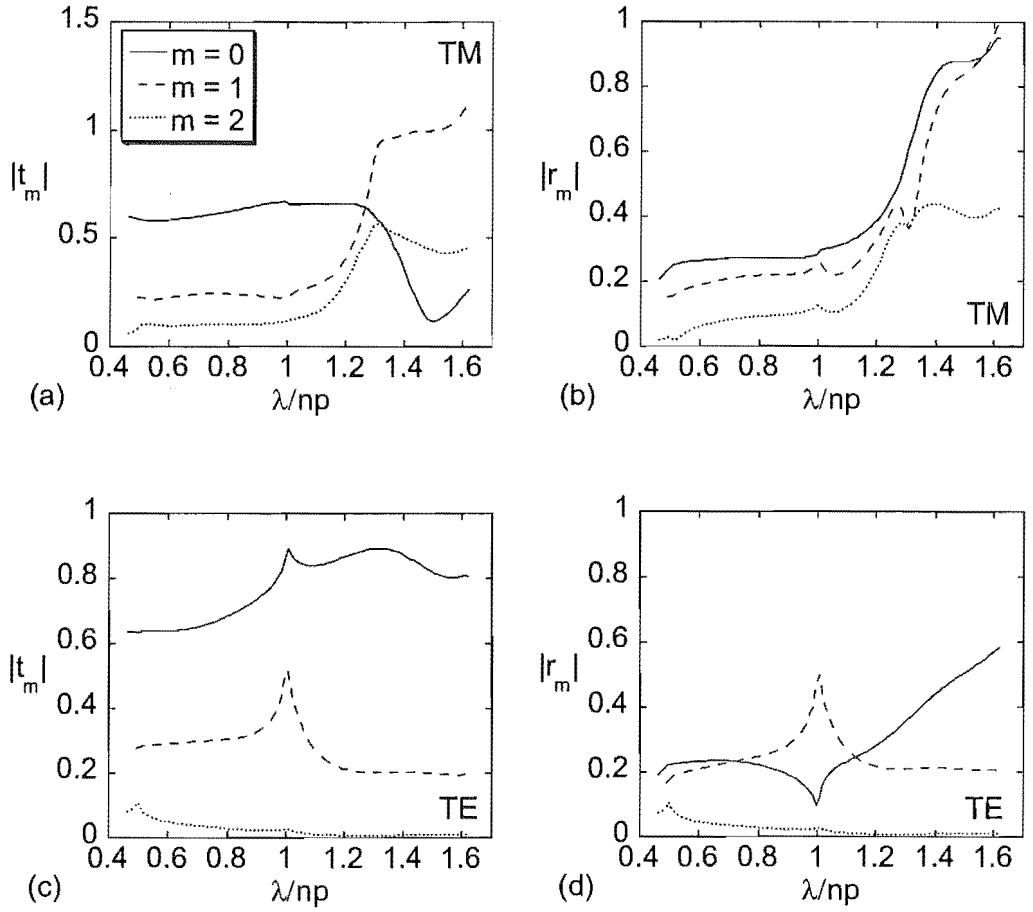


Figure 7.12 Zeroth, first, and second order diffraction (amplitude) coefficients for a gold 270 nm period grating as a function of effective wavelength, normalised by the grating period. The t_m (a) and r_m (b) are plotted for the case of TM illumination and t_m (c) and r_m (d) are plotted for the case of TE illumination.

The diffraction coefficients for Si, plotted in Fig. 7.14 bear little resemblance to either of the metallic gratings. For the TM polarisation, the transmission and reflection coefficients behave similarly with a strong resonance in the t_1 and r_1 components dominating. Both t_1 and r_1 peak in the evanescent regime near $\lambda = 465$ nm. The TE illuminated cases in Fig. 7.14(c) and (d), is rich in resonant behaviour in an extended region around cut-off.

Figure 7.15 illustrates the 2-D intensity maps for the Si grating from Fig. 7.14 at two of the wavelengths of interest (a) at cut-off, and (b) at 465 nm, the wavelength corresponding to the t_1 peak in Fig. 7.14(a). At cut-off there is a very strong interference pattern, more intense than the Cr grating due to higher light transmission. While the period halving of Si appears to be stronger than for the Cr results, using the figure of merit of contrast V to compare the two, Cr performs better with $V = 0.8$ compared to 0.7 for Si due to silicon's overall higher background intensity. In Fig. 7.15(b) high

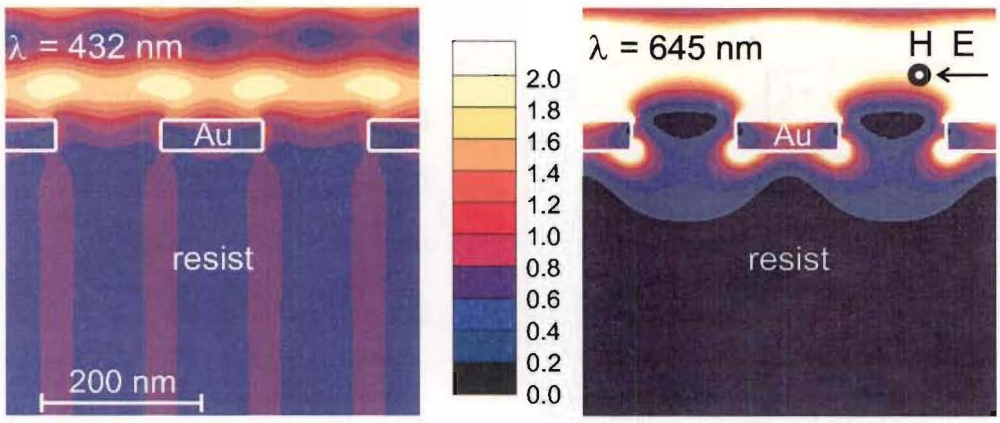


Figure 7.13 Normalised electric field intensity maps of a 270 nm-period Au grating, suspended in photoresist (n = 1.6) and illuminated by TM polarised light (a) $\lambda = 432$ nm (cut-off), (b) $\lambda = 645$ nm. Intensity contours vary linearly from 0 (black) to 2.0 (white) in 10 linear steps.

intensity fields are evident on both sides of the grating, the period halving is still present, but the shadowing of the silicon is poor.

From looking at the transmission coefficients and the intensity field plots for different grating materials it is clear that the choice of material has a big influence on determining the success of EIL. To the author's surprise a dielectric such as silicon can actually perform better than a good conductor such as gold.

Surface plasmon dynamics are very dependent on the material roughness. In fact one way of determining surface roughness is by measurement of the intensities and angular distributions of light [149]. This suggests that the EIL intensity patterns may be strongly dependent on the surface quality of the grating, shifting or altering the magnitude of resonances for the cut-off transmission and reflection coefficients. Factors such as surface profile warrant a more complete study of EIL, however, this is beyond the scope of this thesis.

7.4.2 Effect of Absorbing Resist

Another potentially important issue for EIL is the effect of absorption in the resist. Some amount of loss must be present to allow the photochemical reaction to take place that characterises a photoresist during exposure. The absorption causes a decay in intensity, which as well as increasing the natural decay rate in the transmission direction, also reduces the intensity of the surface waves travelling along the grating that induce the interference pattern. This is expected to adversely affect the depth of field. To evaluate how significant the effect is likely to be, some loss was added to the photoresist in the models, equivalent to a commercial photoresist. Figure 7.16

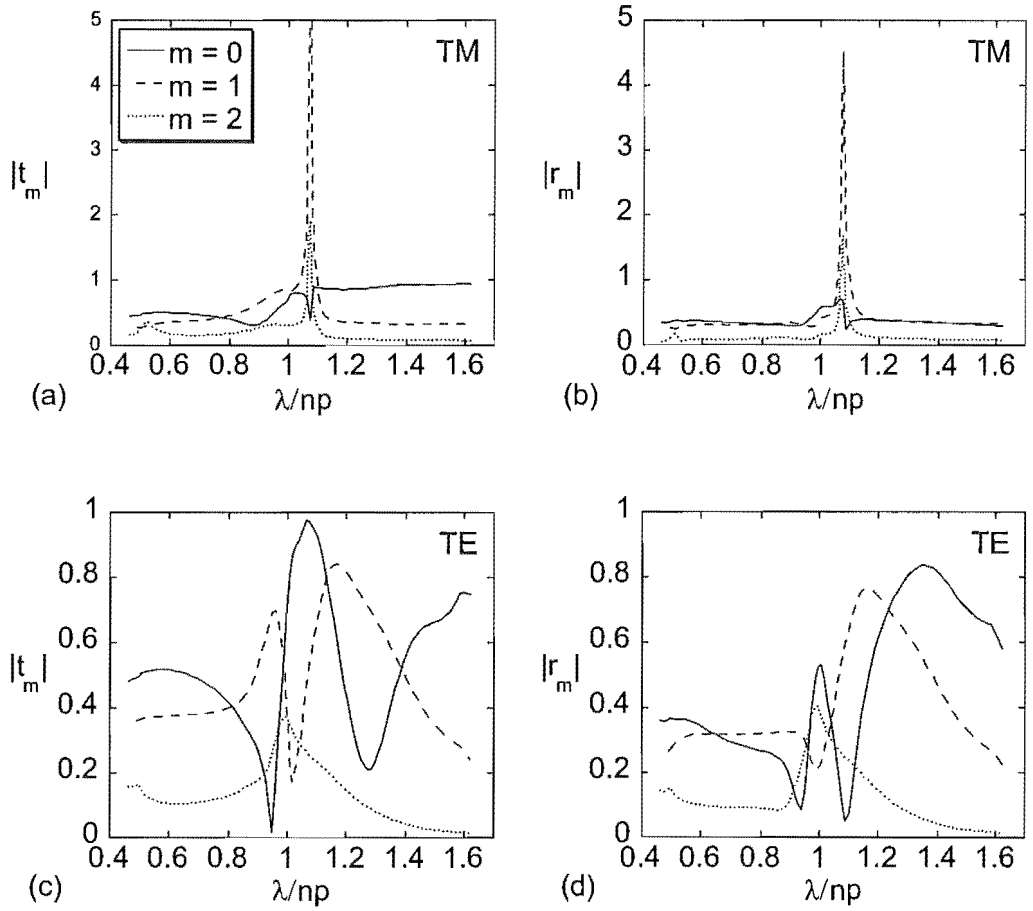


Figure 7.14 Zeroth, first, and second order diffraction (amplitude) coefficients for a silicon 270 nm period grating as a function of effective wavelength, normalised by the grating period. The t_m (a) and r_m (b) are plotted for the case of TM illumination and t_m (c) and r_m (d) are plotted for the case of TE illumination.

compares the field intensities for the resonantly enhanced model described earlier with illumination at 454 nm with the Cr grating suspended in resist with four different absorptions. Figure 7.16(a) represents the lossless case shown earlier (n = 1.6), while (b) and (c) have values for the extinction coefficient (n'') taken from Clariant resist property tables for the AZ1500 series resist at 436 nm [150] for an unbleached resist (n = 1.6 + i0.0227) and a bleached resist (n = 1.6 + i0.0003) respectively. During the exposure process, the resist bleaches as the photoactive compound is used up and the resist becomes more transparent. While some evidence of damping can be discerned from the intensity maps caused by the addition of loss, the effect appears minor at this level of absorption. Figure 7.16(d) shows the resultant intensities when a strongly absorbing resist is modelled – (n =1.6 + i0.1338). This equates to an absorption for the electric field (α_E of 3.7037 μm^{-1}) which equates to an absorption depth of 270 nm, *i.e.*, one period.

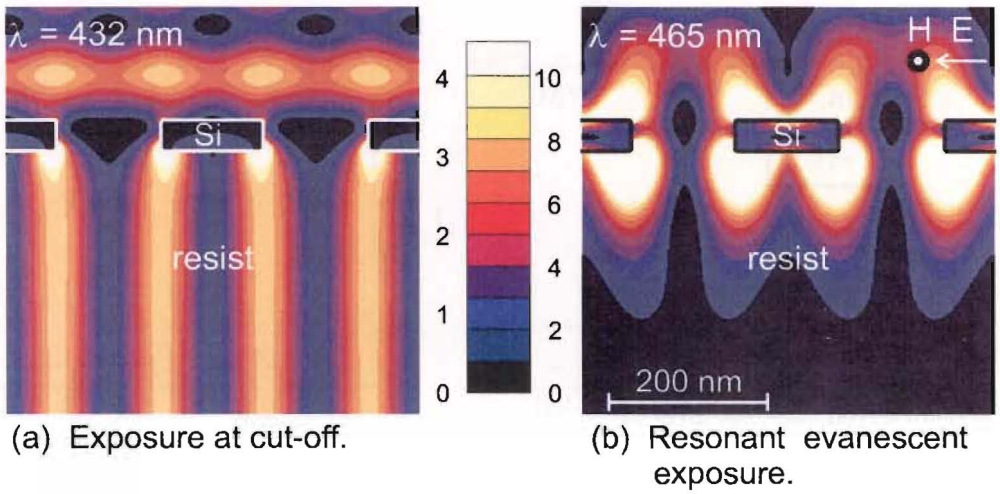


Figure 7.15 Normalised electric field intensity maps of a 270 nm period Si grating, suspended in photoresist (n = 1.6) and illuminated by TM polarised light (a) $\lambda = 432$ nm (cut-off), (b) $\lambda = 465$ nm. Contours vary linearly from 0 (black) to 2.0 (white) in 10 linear steps.

Line-plots of the normalised intensity for the lossless, unbleached, and strongly absorbing resist are compared in Fig. 7.17 at the exit aperture of the grating, and at 10, 50, 100 and 200 nm below the grating. From the line-plots in Fig. 7.17 it can be seen that the intensity is reduced by the addition of loss to the photoresist. For the unbleached case, the loss has a significant effect but exposure should still be feasible with some reduction in the depth of field. The strongly absorbing resist shows very little contrast between the intensity null in the aperture compared with the intensity at the edges after 10 nm below the grating.

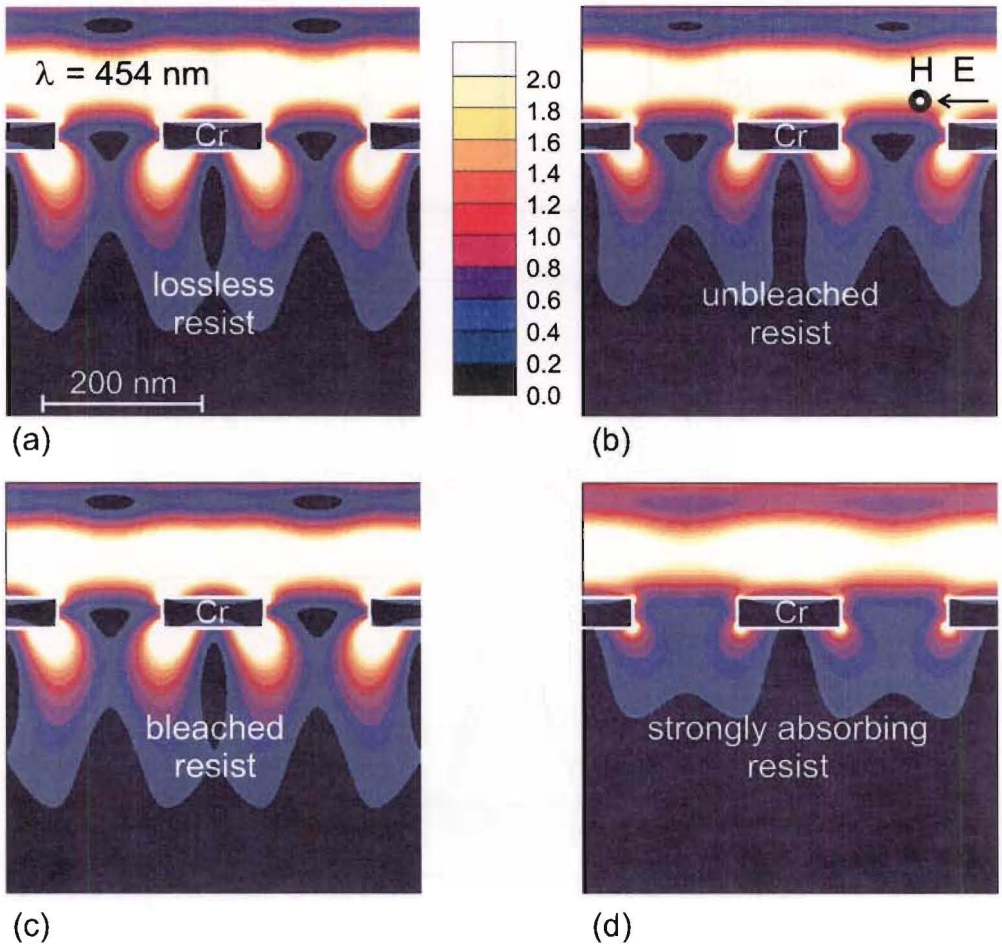


Figure 7.16 Normalised electric field intensity maps of a 270 nm-period Cr grating, suspended in photoresist and illuminated by TM polarised light at resonance ($\lambda = 454$ nm) (a) lossless photoresist ($n = 1.6$), (b) unbleached photoresist ($n = 1.6 + i0.0227$, (c) bleached photoresist ($n = 1.6 + i0.0003$), (d) a strongly absorbing resist ($n = 1.6 + i0.1285$). Intensity contours vary linearly from 0 (black) to 2.0 (white) in 10 linear steps.

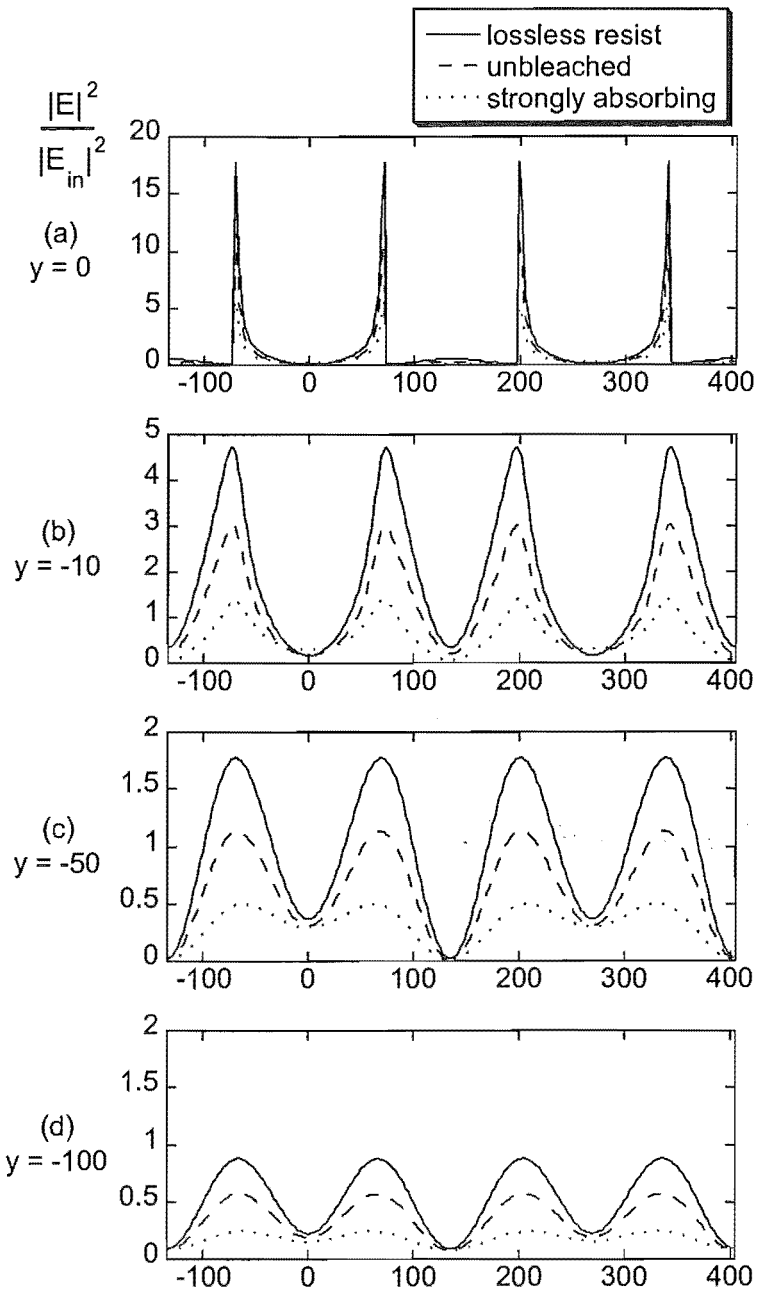


Figure 7.17 Normalised electric field intensity line plots from Fig. 7.16 comparing the intensity when the grating is surrounded by a lossless photoresist ($n = 1.6$), unbleached photoresist ($n = 1.6 + i0.0227$, and a strongly absorbing resist ($n = 1.6 + i0.1285$) at various depths (a) $y = 0$ nm (the exit aperture), (b) $y = -10$ nm, (c) $y = -50$ nm, (d) $y = -100$ nm, (e) $y = -200$ nm.

7.5 SUMMARY

The nature of a grating exposure changes markedly from large period to small period gratings, where the orders that provide contrast change from propagating to evanescent. More complex structure is evident for exposures of large period gratings in the far-field, due to the rich number of diffracted propagating orders that interact, creating effects such as Talbot multiplicative period division which is observed at specific distances. Small period gratings are simple in comparison, their far-fields consist of only a single zeroth propagating order and appear like polarisers. The ranges of the evanescent orders are too small to create interesting interference patterns.

In the transition region, where the wavelength approaches cut-off of the ± 1 diffracted orders, the picture is different. In this region, the ranges of the diffracted orders are sufficient for interference to occur, and are susceptible to redistributions of energy that are common near cut-off thresholds. A period halving is evident in the TM polarisation as a result of interfering surface waves comprising the ± 1 diffracted orders. Additionally the TE polarisation produces an interesting checker-board high intensity pattern. As long as a thickness of resist is selected that is less than one period of the checker-board, an acceptable exposure for standard ENFOL will be obtained for the TE polarisation. However, for the TM polarisation, immediately below the grating, there is a period halving that will be undesirable for standard ENFOL.

To reduce the impact of near-field interference effects, such as that arising with TM light at cut-off, exposures should be made with unpolarised, broadband light that is incoherent and preferably incident at a range of angles. However, specialist applications may benefit from optimisation of the interference to obtain high contrast period halving. To do this, good polarisation control is required, as well accurate information regarding the refractive indices of the photoresist and wavelength control. Some advantages were found for operation in the evanescent regime, at a resonant peak of the $t_{\pm 1}$ coefficients, close to a t_0 minimum for a Cr grating. Operating at this resonance resulted in a higher intensity interference pattern with better contrast over a limited, but workable distance from the grating. The resulting trade-off between improved intensities and depth of field indicates that the ideal would be a large resonant peak for $t_{\pm 1}$ as close to cut-off as possible to extend the depth of field. The interference effect shown here can be likened to the interference lithography described earlier for propagating waves. This is performed with a phase grating, with the depth of the grating selected to suppress the zeroth order. However for the EIL case, a resonance of an amplitude grating is taken advantage of and the interference is evanescent in the transmission direction.

The grating material is shown to be an important factor dictating the quality of the interference pattern. Of the material properties compared, Cr proved to have the most favourable qualities at resonance, although surprisingly silicon seemed to have a far enhanced period halving at the first order cut-off. While an exhaustive set of materials

was not investigated, it is clear that this is an important parameter to optimise. The strong reliance of the exposure on material properties suggests that variations in the deposition of the grating films and surface properties may also have a large effect on results. However, EIL appears a feasible proposition worthy of some experimental effort and further simulation work, particularly if a grating material with improved material properties could be found.

Chapter 8

WIRE-GRID POLARISER

8.1 INTRODUCTION

A wire-grid linear polariser was fabricated using ENFOL to serve as a proof of concept device. The principle of the wire-grid polariser is illustrated in Fig. 8.1. It consists of a one-dimensional grating of conducting wires with a period p and grating vector k_g. When illuminated by unpolarised light of a wavelength significantly greater than the grating period (ideally $> 10p$), the component of the light having its electric field parallel to the grating vector (k_g) is preferentially transmitted. The component of light with its electric field perpendicular to the grating vector (k_g) is preferentially reflected, and there are some losses in both polarisations owing to absorption in the grating material. This polarisation effect was observed in the simulation results in Section 6.5.

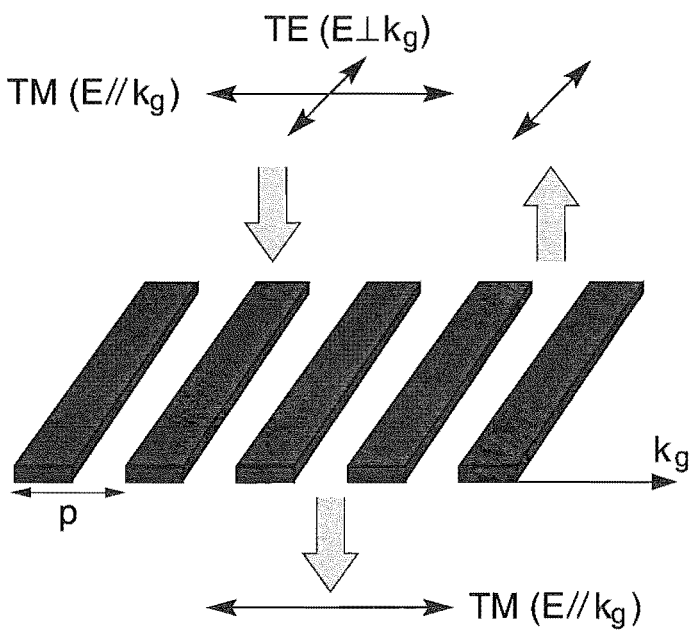


Figure 8.1 Schematic illustrating the principle of a wire-grid polariser.

Wire-grid polarisers have been generally limited to wavelengths in the infra-red range and longer owing to the difficulties in making wire-grid polarisers with small grating periods. Fabrication of wire-grid polarisers has been reported for the visible range using electron beam lithography [151] but its slow serial nature prohibits economic manufacture of large quantities. With the recent advance of low-cost, high resolution fabrication techniques such as nanoimprint and ENFOL, realisable manufacturable solutions look promising [152].

The wire-grid polariser is an ideal device for demonstrating the potential of the ENFOL technique. It is a relatively simple device, that benefits from high resolution. The simplicity of the polariser stems from the fact that only a single mask is required, avoiding the need for multi-level alignment. While the possibility of multi-level alignment is not ruled out with ENFOL, it is a challenging issue that has not been addressed in this work as yet. Line-edge roughness present in the pattern is also not a critical problem for polarisers. The performance of the polariser is determined by the macroscopic properties of the device as opposed to being critically sensitive to line irregularities, as ICs are. However, the better the pattern quality, the better the expected performance of the polariser.

The wire-grid polariser has potential applications as a device for integrated optics applications in communications and other related areas. It provides a compact substitute compared to more bulky 3-D birefringent crystal-based polarisers, and may be more easily integratable than polymer polarisers [151]. The availability of a suitable low-cost, high resolution manufacturing technique would make the wire-grid polariser a desirable alternative for some applications. This chapter describes the fabrication of wire-grid polarisers using the ENFOL technique. Measurements of the extinction ratios of the devices are reported for these initial devices and comparisons are made with simulated models.

8.2 FABRICATION PROCEDURE

Polarisers were fabricated onto 200 μm thick glass substrates using an ENFOL exposure and lift-off metallisation with a trilayer resist scheme. The trilayer resist comprised a Brewer XLT anti-reflection coating (ARC), SiO_x and S1805 photoresist thinned 1:4. Details of the processing have been previously described in Section 5.4.4. A 65 nm thick layer of NiCr was thermally deposited to form the grating absorbers. Polarisers with grating periods of 270 nm and field sizes of approximately $60 \times 40\,\mu m^2$ were fabricated. AFM images of two separate fields of the fabricated polarisers are shown in Fig. 8.2. Figure 8.2(a) shows a field that has successfully lifted off. The lines of absorber are quite distinct, but some of the lines appear to be thicker in height in places. This may be due to redeposition of metal from the sidewalls during the evaporation process. Figure 8.2(b) shows a field that had unsuccessful lift-off. The high sections of the

topography are metal on top of the trilayer resist. Some of the lines are skewed in places, and in a few areas, gaps suggest there has been partial lift-off. Polariser fields that did not lift-off were also measured and are referred to as double-layer grating polarisers in the remaining section.

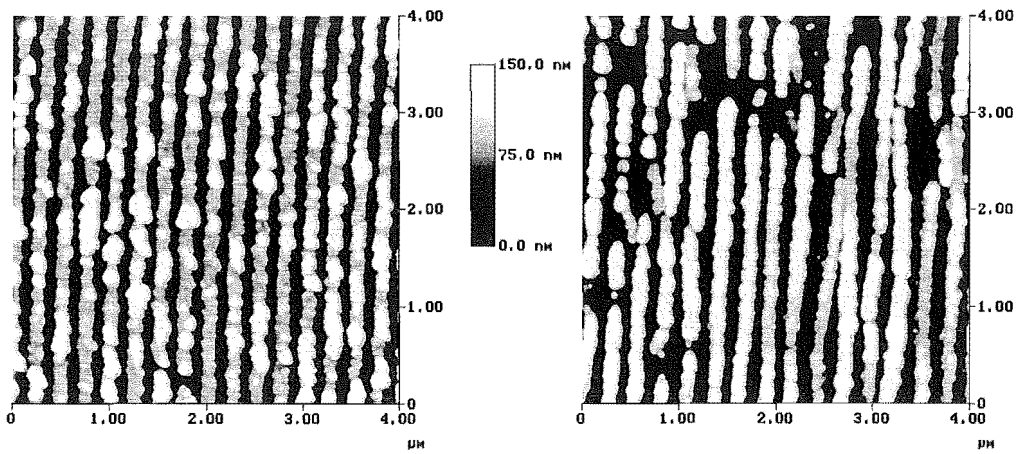


Figure 8.2 AFM images of two separate polariser fields, (a) successful lift-off (single grating), (b) unsuccessful lift-off (double grating). Both fields have a grating period of 270 nm, and 65 nm of NiCr deposited.

8.3 POLARISATION MEASUREMENTS

To assess the performance of the wire-grid polarisers, measurements were made of their transmitted extinction ratios T_{TM}/T_{TE}, where T_{TM} is the transmitted intensity of the TM polarised light, and T_{TE} the transmitted intensity of the TE polarised light. Measurements were made with an optical microscope, the Olympus BX60, operating in transmission mode with a digital camera add-on, the Olympus DP10. Figure 8.3 illustrates the setup.

The illumination aperture was made as small as possible by closing down the field and condenser iris diaphragms to minimise scattering. The aperture could not be too small though as adequate light was needed for the camera exposure. All of the light source's neutral density filters were de-activated to provide as intense a broad-band source as possible. The wire-grid polariser to be measured was aligned in the centre of the light source. The image is magnified with the maximum objective lens available ($\times 100$) and focused. All optical devices in the light path incidental to the measurement were removed. This was particularly important in the case of the bright-field and dark-field mirror cubes that are used in reflection mode microscopy to direct the light onto the sample, as these added some additional polarisation. The light, following transmission through the sample, is then passed through an adjustable polariser (the U-AN360).

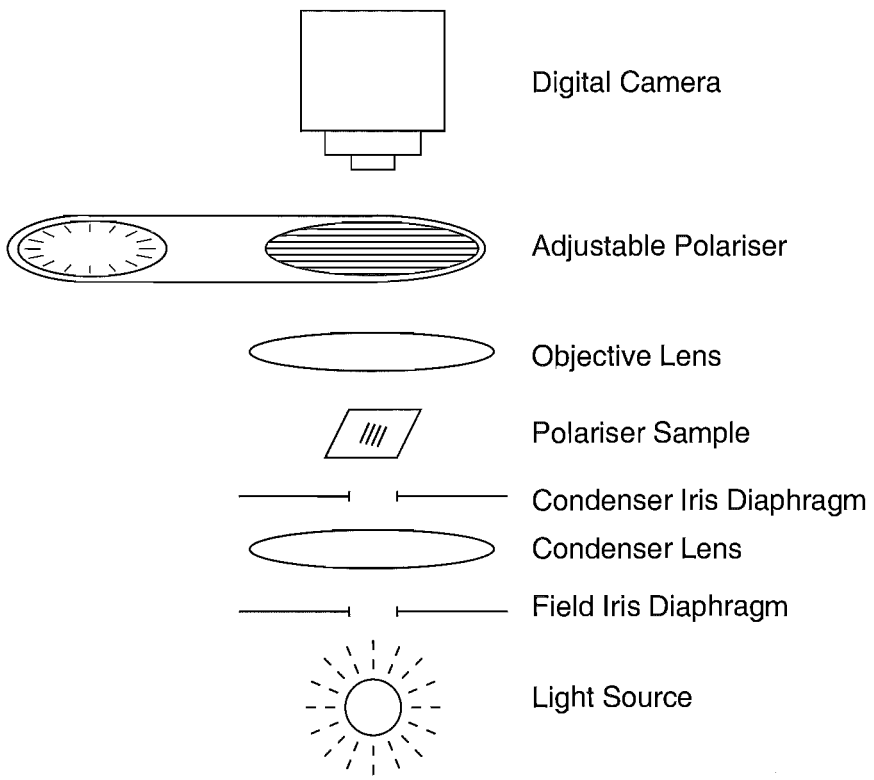


Figure 8.3 Setup for measuring the polariser's transmission extinction ratio.

The angle of this polariser can be stepped in 2° increments. The doubly polarised light, initially polarised by the polariser sample and lastly from the U-AN360, is then captured by the digital camera. The camera contains RGB primary colour on-chip filters with a 1 410 000 pixel array and 8 bits per pixel.

The extinction coefficient of the adjustable U-AN360 polariser was assessed in conjunction with an additional UV polariser. The maximum extinction ratio possible was obtained between crossed and uncrossed polarisers with the 8 bit measurement setup, indicating that the adjustable polariser would not limit the measurements. Prior to taking measurements, the polariser field was located and the adjustable polariser rotated to give maximum transmittance. The camera was then set to take a photograph and the exposure locked at that position. The polariser was then adjusted for minimum transmittance and a photograph taken every 10°. The images were saved as JPEGs, consistent with the internal format of the camera with an underlying resolution of 1024 x 1280 pixels and three colour channels. As the polariser fields do not take up the entire field of view of the microscope/camera, the section of the image containing the polariser was identified for analysis. The mean intensity over the polariser field for each step in angle of the adjustable polariser was calculated for each colour channel. This setup was not as accurate as one based on a spectrophotometer, however it was

adequate for obtaining initial polariser measurements. The measurement setup could be established quickly, and importantly, as the polariser fields were quite small, the polarisers could be aligned with the light source easily using the built in microscope for magnification.

Figure 8.4 illustrates the transmitted intensities for two polarisers as the angle ϕ is rotated between 0° and 180°. The results plotted in Fig. 8.4(a) are for a 270 nm-period NiCr grating that has lifted off, corresponding to the AFM image in Fig. 8.2(a). The results plotted in Fig. 8.4(b) are for a 270 nm-period NiCr grating that has not lifted off, corresponding to the AFM image in Fig. 8.2(b). The transmitted values are normalised to those of the background light through 65 nm of NiCr on glass. The polarising nature

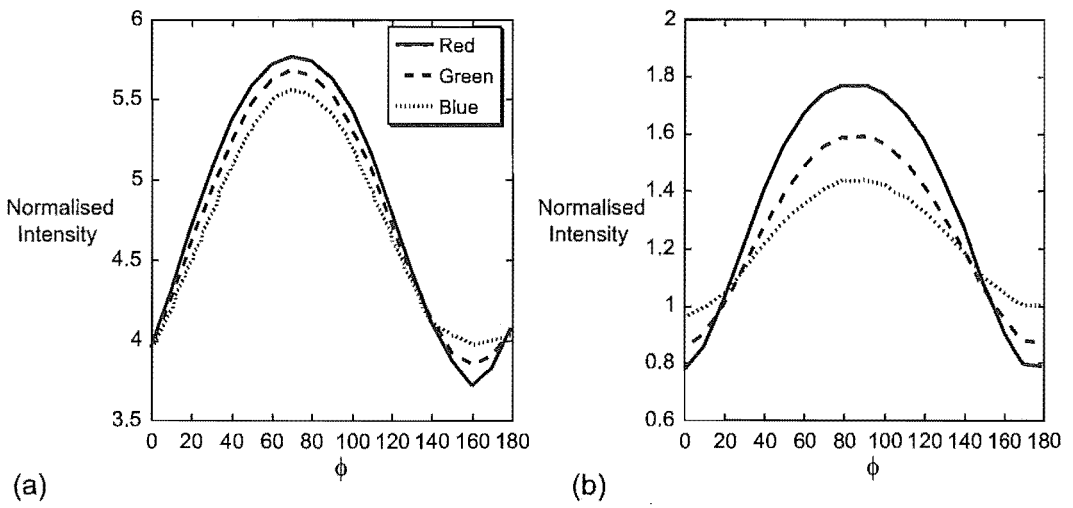


Figure 8.4 Measured transmitted intensity versus angle of the adjustable polariser for fabricated 270 nm-period NiCr on glass polarisers. (a) single layer 65 nm thick NiCr grating (lifted-off), (b) double layer grating 65 nm NiCr (no lift-off).

of the gratings at visible wavelengths can be clearly seen. As the adjustable polariser is rotated through 180°, the transmitted intensity reaches a maximum as the polarisers are aligned, then goes through a minimum as the polarisers cross at right angles. Note that the transmission peaks in Fig. 8.4(a) and (b) are not aligned owing to the different initial starting angle of the adjustable polariser relative to the polariser field. The red (solid) curve, shows the greatest change in transmitted intensity, followed by the green (dashed) curve and finally the blue (dotted) curve. As expected, the fabricated grating is a more effective polariser at longer wavelengths due to the larger ratio of the wavelength to grating period. The transmitted intensity is significantly higher than the background intensity for the polariser at all angles for the lifted-off polariser (Fig. 8.4(a)). In comparison, the normalised intensities for the polariser without lift-off (Fig. 8.4(b)), are a factor of three smaller, but otherwise show the same trends, with the best polarisation achieved at longer wavelengths. It should be noted that when the unlifted-off polariser is at right angles to the adjustable polariser (equating to minimum

transmission), the intensity of the transmitted light is less than through the 65nm-thick NiCr background.

Table 8.1 summarises the extinction ratios for the polarisers plotted in Fig. 8.4. The polarisers that did not lift-off have higher extinction ratios, due to what is effectively a double grating, offset by half the grating period. Note the improved extinction ratio comes at the cost of reduced intensity. While the extinction ratios from either the single or double gratings are not high, the principle is shown.

Table 8.1 Extinction ratios measured from polarisers with and without lift-off.

Wavelength	Lift-off T_{TM}/T_{TE}	No Lift-off T_{TM}/T_{TE}
Red	1.6	2.3
Green	1.5	1.9
Blue	1.4	1.5

8.4 POLARISER SIMULATION RESULTS

GSolver simulations of polariser structures were performed to compare the performance of the fabricated polarisers. GSolver is a commercially available electromagnetic simulation technique based on the rigorous coupled wave analysis technique described in Section 6.2.3. It has also been used previously for validation of some of the MMP simulation work described in Section 6.7. GSolver was used in preference to MMP as the models are simpler and less time consuming to formulate. For the polariser simulations, only the far-field transmission coefficients are of interest as opposed to the near-field information required in Chapter 6 and Chapter 7.

Simulations at a representative wavelength in the middle of the red, green and blue visible bands were performed; these wavelengths are specified in Table 8.2. The GSolver models are illustrated in Fig. 8.5. Figure 8.5(a) represents a lifted-off polariser, constructed with seven simulation layers — a silica substrate ($n = 1.5$), an air superstrate ($n = 1$), and a five-layer 65 nm thick Cr grating interspersed with air ($n = 1$). The Cr has a complex refractive index n that is wavelength dependent. The values of n for the simulated wavelengths were interpolated from published tabulated values [94] and the values used are given in Table 8.2.

A second model that represents a polariser prior to lift-off is represented by a double-layer grating in Fig. 8.5(b). This model comprises thirteen simulation layers, with a silica substrate, air superstrate, and two five-layer 65 nm thick Cr gratings separated by a 100 nm thick ARC/air grating layer. The bottom layer Cr grating is interspersed with ARC ($n = 1.74 + i0.34$), while the top grating is interspersed with

Table 8.2 Wavelengths used for the polariser simulations using GSolver and the refractive indices for Cr corresponding to these wavelengths.

Colour Band	Wavelength (nm)	Chrome n
Red	680	3.76 + i4.37
Green	538	2.96 + i4.44
Blue	473	2.18 + i4.33

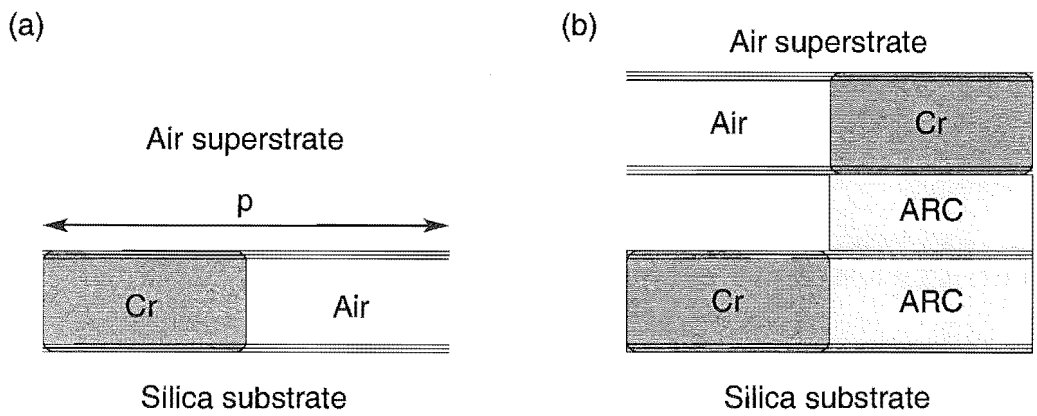


Figure 8.5 Polariser models constructed for GSolver simulations showing the different simulation layers constructed for (a) a single-layer Cr grating and (b) a double-layer Cr grating.

air. The Cr gratings have each been split into five individual simulation layers to provide some stepping of the Cr conductor corners, avoiding an infinitely sharp corner.

Simulation results for the single- and double-layer grating polarisers described above are plotted in Fig. 8.6. The zeroth-transmitted order (T_0) is plotted as a function of the angle of the incident electric field, where $\phi = 0°$ equates to the electric field direction perpendicular to the grating vector k_g, and $\phi = 90°$ equates to the electric field direction parallel to the grating vector. Table 8.3 summarises the simulation results in terms of the transmission extinction ratios for the single- and double-layer gratings.

Table 8.3 Simulated transmission coefficients and extinction ratios for single-layer and double-layer gratings of 65 nm-thick Cr.

Illumination Source	Single Layer Grating			Double Layer Grating		
	T_{TM}	T_{TE}	T_{TM}/T_{TE}	T_{TM}	T_{TE}	T_{TM}/T_{TE}
Red	0.54	3.3E-2	16.4	0.115	7.8E-4	148.5
Green	0.41	5.3E-2	7.7	0.060	1.7E-3	35.6
Blue	0.30	8.9E-2	3.4	0.039	3.7E-3	10.6

The extinction ratio (T_{TM}/T_{TE}) improves at longer wavelengths for both the single-

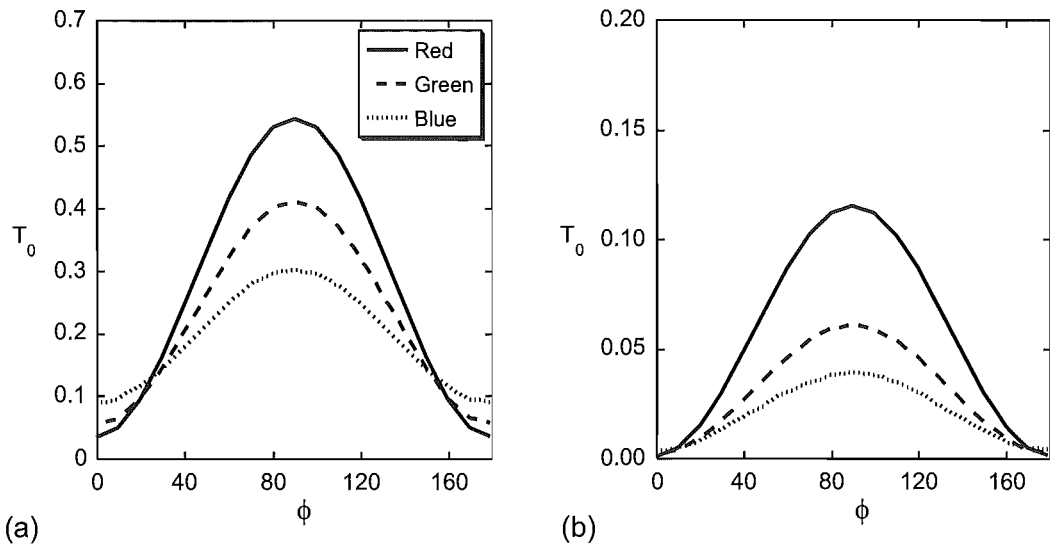


Figure 8.6 Simulated transmission coefficients for a 65 nm-thick Cr grating illuminated with TE light (0°) to TM light (90°) for (a) a single-layer grating and (b) a double-layer grating.

and double-layer gratings as is expected as the λ/p ratio gets larger. This is consistent with experimental measurements. The simulations also show that the double-layer grating is a better polariser than the single-layer grating with an extinction ratio nine times higher in the red. However, the improvement comes at the cost of a reduction in the overall transmission intensity. If this can be tolerated, the double-layer grating polariser provides improved performance with the advantage of simplifying the processing, by removal of the lift-off step.

The experimental and simulated results agree qualitatively, however the extinction coefficients are significantly lower than the experimental results. There are many factors that contribute to the differences, including the quality of the polarisers, differences between the simulated model to the experimental system and deficiencies in the measurement setup. As can be seen from the the AFM images of the polarisers shown in Fig. 8.2, there is a certain amount of line-edge roughness that may affect the results. The duty cycle of the lifted-off grating also appears to be less than the 50% simulated. From the AFM image of the grating that did not lift-off (Fig. 8.2(b)), it is not surprising that measured extinction ratios are lower, as the top NiCr grating appears to be damaged, presumably during the solvent soaking stage. It is expected that the quality of the double-layer grating, and hence its polarising performance would be improved by omitting the solvent soak step.

There are a number of differences between the fabricated polarisers and simulated models. Firstly, as the material properties of NiCr were unavailable, Cr was modelled as the material for the grating absorber. Different material properties can have significant effects at optical frequencies where the skin depth is appreciable. Secondly,

only one representative frequency is modelled for the red, green and blue, whereas the experimental measurements were broad-band.

The measurement system will contain some scatter in the system; this will predominantly contribute to error in the measurement of the minimum transmissions, obtained when the polarisers are crossed. The source is also uncollimated, which may contribute to the scatter. The camera stores images in a JPEG format which incorporates some data compression that will alter the raw data, introducing an additional unknown factor. While this is not ideal, it is believed to be only a minor contributing factor to errors in the measurement. When taking the photographs it is important that the exposure time (which is fixed for all the measurements) does not saturate the pixels at any point within the measurement. Too long an exposure saturates the pixels and results in an underestimate of the maximum transmitted light. Conversely with not enough light the minimum transmission can be underestimated. Both conditions result in an underestimation of T_{TM}/T_{TE}. Obtaining the correct exposure was a matter of trial and error, and could only be properly determined once the images had been taken by manual assessment of the data.

These experimental results represent polariser extinction ratios for an unoptimised wire-grid polariser. Further work is expected to significantly improve the polariser performance. Doumuki *et al.* reported improved performance of fabricated polarisation selective photodiodes by optimising the grating duty cycle and absorber thickness [153]. They were able to obtain extinction ratios of 30 with grid periods only slightly smaller than the illumination wavelength ($p = 600$ nm, $\lambda = 715$ nm).

The most immediately obvious way to improve the ENFOL fabricated polarisers would be to deposit a thicker absorbing layer. Simulations increasing the absorber layer from 65 nm to 100 nm were performed and the results are plotted in Fig. 8.7. Table 8.4 summarises the simulated transmission intensities for each of the linear polarisations and the resultant extinction ratios for the 100 nm-thick Cr gratings.

Table 8.4 Simulated transmission intensities and extinction ratios for 100 nm-thick Cr single-layer and double-layer gratings.

Illumination Source	Single Layer Grating			Double Layer Grating		
	T_{TM}	T_{TE}	T_{TM}/T_{TE}	T_{TM}	T_{TE}	T_{TM}/T_{TE}
Red	0.52	0.009	55	0.075	5.1E-5	1154
Green	0.46	0.018	25	0.077	2.6E-4	296
Blue	0.40	0.034	12	0.021	7.0E-4	30

From the simulations of the 100 nm-thick polarisers it can be seen that even a small increase in the absorber thickness from 65 nm to 100 nm results in a better than three-fold improvement in T_{TM}/T_{TE} for the single-layer grating in the red. For the double-layer grating the improvement is even larger with a better than seven-fold improvement in T_{TM}/T_{TE}, although low transmission intensities (7.5%) are predicted.

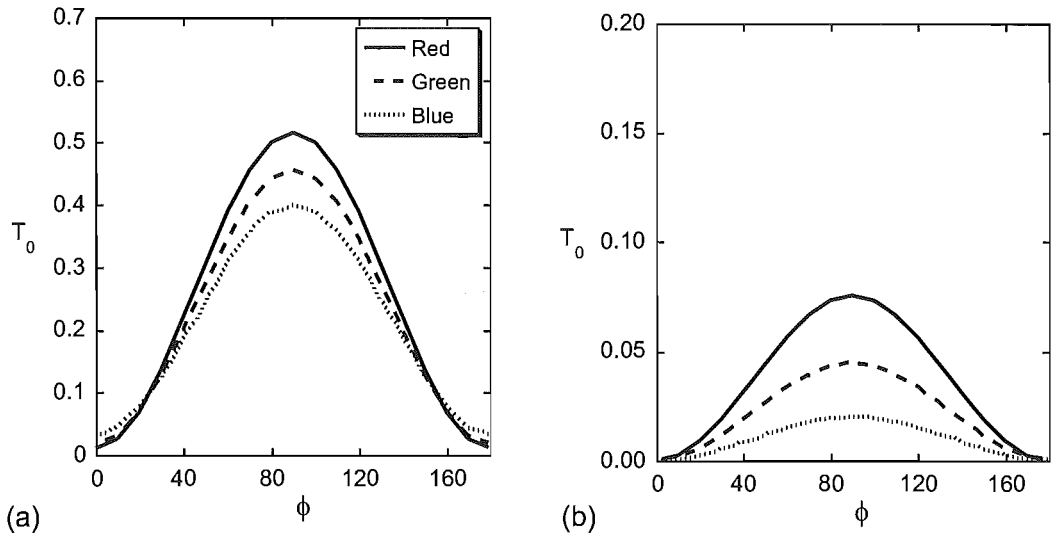


Figure 8.7 Simulated transmission coefficients for a 100 nm-thick Cr grating illuminated with TE light (0°) to TM light (90°) for (a) a single-layer Cr grating and (b) a double-layer Cr grating.

Double-layer wire-grid polarisers have been reported in the literature to exhibit high extinction ratios when operating in reflection mode [152]. Reflection extinction ratios of over 200 were obtained as opposed to transmission extinction ratios of close to one. Operation in reflection mode also has the advantage of high intensities compared to transmission mode.

Another parameter that could be optimised to improve the transmission extinction ratios is the choice of metal absorber. A higher conductivity metal is expected to give improved results; currently only NiCr has been used as it has demonstrated good lift-off for small features. For the double-layer gratings, the intermediate layer thickness and material are additional parameters that can be optimised. Unfortunately the lift-off resist, in this case the XLT ARC, has conflicting requirements. For the ENFOL exposure, a highly absorbing material is desirable to prevent reflections, while for the final polariser device, low absorption is preferred to prevent unwanted attenuation of the transmitted intensity. A low refractive index for the intermediate layer is also desirable, so that the effective wavelength appears as large as possible to improve the grating's polarising properties.

8.5 SUMMARY

Wire-grid polarisers have been fabricated using an ENFOL exposure and additive pattern transfer using a trilayer resist process. The optical performance of two different polariser variants was evaluated. Firstly, a conventional single-layer grating polariser, and secondly, a double-layer grating structure corresponding to a grating without lift-off. The double-layer grating polariser demonstrated an extinction ratio of 2.3 in the

red, a higher extinction ratio than for the equivalent single-layer grating polariser. An advantage of the double-layer grating polariser is that is it easier to fabricate, however, the intensity of the transmitted polarisation is smaller, in this example by a factor of three.

Simulations were performed using rigorous coupled wave analysis, modelling similar structures to that fabricated and measured. Qualitatively consistent results were obtained, with predictions of improved extinction ratios for the double-layer grating polarisers. The trend of better polarisation at longer wavelengths was also confirmed. Quantitatively, the simulations predicted higher extinction ratios than were measured, a factor of approximately 10 and 60 higher for the single- and double-layer gratings respectively in the red. Differences between the experimental and simulation results can be accounted for by the quality of the fabricated gratings, particularly the double grating layer, as well as differences in the simulated model and scatter in the measurement system.

While polarisers with only relatively small polarisation extinction ratios have been demonstrated, large improvements are expected with further optimisation of grating parameters even without decreasing the grating period. The single-layer grating would benefit from the deposition of a thicker absorber, however simulations with thicker absorbers for the double layer grating predict very low intensity for the preferentially transmitted polarisation.

The fabrication of the wire-grid polariser demonstrates that ENFOL is a viable technique for fabrication of devices. ENFOL's capability of high resolution at low cost is a particularly desirable characteristic that makes it ideal for niche fabrication areas. Optical devices such as polarisers and anti-reflection coatings are particularly suitable for techniques such as ENFOL, that are capable of delivering the sub-diffraction limited resolution that is required of these devices.

Chapter 9

CONCLUSIONS AND FUTURE WORK

In this thesis a study of an evanescent near-field optical lithography (ENFOL) technique has been presented, the result of a combination of experimentation and simulation. Membrane masks capable of testing the resolution of ENFOL were fabricated, that are an essential pre-requisite for ENFOL studies. Characteristics of an ENFOL exposure were investigated, and an additive pattern transfer technique developed, which was the main goal of the experimental work. While additive pattern transfer techniques have been developed previously, the challenge was to develop a process compatible with the ENFOL's constraining requirement of thin resists. The experimental work culminated in the fabrication of a sub-wavelength device using ENFOL, demonstrating ENFOL's potential for fabricating devices with high resolution at low cost.

Results from rigorous electromagnetic simulations of grating structures have shown encouraging prospects for the attainable resolution of near-field techniques such as ENFOL. Valuable feedback for the experimental set-up was obtained for optimising the exposure of sub-wavelength gratings and possible pitfalls have been identified.

The contributions to ENFOL presented in this thesis are just a starting point for a technique currently in its infancy. Potential exists for improved resolution with greater reliability, with progress in mask and resist technology in particular. In the remainder of this chapter the results presented in this thesis are summarised and suggestions for future directions indicated.

9.1 EXPERIMENTAL

The experimental achievements of this research include the fabrication of membrane masks for ENFOL experiments. Silicon nitride membrane masks were fabricated that were conformable, consistent with ENFOL's requirement for uniformly intimate contact between the mask and substrate during an exposure. Mask patterns down to minimum feature sizes of 70 nm were defined using electron beam lithography, demonstrating resolution close to the limit of our electron beam system. While the membrane masks fabricated were successfully used for demonstrating the potential for high reso-

lution with ENFOL, future studies would benefit from improvements in both the mask materials and the mask fabrication process.

Mask profile issues became apparent during the course of this work. The absorber patterns on the mask appeared to be thickened at the edges by $\sim$ 10nm during the metal deposition and lift-off process. Any factors that act to increase the gap in the ENFOL process will have a detrimental effect on resolution and should be minimised. Future generation ENFOL masks may benefit from a subtractive pattern transfer technique to improve the profile of the mask absorbers. Tungsten has been proposed as a possible alternative for the mask absorber, owing to its etchability and suitable electromagnetic properties.

Mask robustness is another serious issue for ENFOL. The robustness of membranes was inconsistent, with some masks appearing very hardy, and others extremely fragile. Some initial attempts to improve the taper of the membrane support to the membrane edge by isotropic etching were made, but a more comprehensive evaluation is required to assess the success of this approach. A different mask substrate material may be beneficial. While silicon nitride has some favourable properties, such as low stress and low sticking problems, it is limited to exposures at blue wavelengths or longer due to the high absorbance below these wavelengths. Silica glass is an alternative mask material that was used by Goodberlet [47] that is transparent down to deep-UV wavelengths.

Pattern placement error is an issue that has not been addressed for ENFOL as yet. Flexible, conformable masks can introduce in-plane distortions that shift the relative positions of features during the exposure. These distortions will also compromise multi-level alignment. Planarisation of the substrate topography can alleviate these problems to some extent, as may masks such as the embedded-amplitude mask proposed by Goodberlet [47] where a flat mask surface is obtained by embedding the absorber into the mask substrate. Mechanical simulations of the conformability of various mask substrate materials at different thicknesses may be helpful for optimising the need for conformability, pattern accuracy and mask robustness.

The trilayer resist system developed for the lift-off metallisation process was successful at patterning the smallest resolved photoresist images, 270 nm period gratings. However, pattern transfer of yet smaller features, or the deposition of thicker metal layers, may require further improvements in the trilayer resist processing. Side-wall profiles of the trilayer resist may benefit from the addition of a side-wall inhibitor such as methane for the bottom layer etch [154]. If thicker metal layers needed to be deposited, a barrier layer material with higher etch selectivity may be required so that thicker bottom layers can be etched. A potential candidate for the barrier layer is Ta_2O_5 which has demonstrated improved selectivity compared to silicon dioxide [90].

Wire-grid polarisers were fabricated onto silica substrates using an ENFOL exposure followed by the additive pattern transfer scheme developed in this work. Two

variations of wire-grid polarisers were compared, with a single and a double layer grating structure. The measured extinction coefficients of the polarisers were not high, an extinction ratio of 2.3 was obtained at red wavelengths. However, the principle of operation was shown, and the potential of ENFOL for device fabrication demonstrated. Optimisation of the double-layer grating may be worthwhile, as it is simpler to fabricate and demonstrated superior extinction ratios compared to the standard single-layer grating polariser. The polariser performance is expected to improve with the deposition of a thicker metal layer, although for the double-layer grating this is at the cost of low transmitted signal power. The integration of polarisers onto a device such as a photodiode would be an interesting demonstration of the ENFOL technique.

9.2 SIMULATIONS

The key finding from the simulation results reported in this thesis is the potential resolution of grating structures down to $\lambda/20$. This prediction of resolution is significant from two viewpoints. Firstly, resolution of $\lambda/20$ periods is well below (by a factor of ten) that of the conventional diffraction limit defined for optical projection based systems. Resolutions in this order have been previously reported for near-field scanning microscopy applications, but this is the first time such a claim has been made for a parallel optical lithography based system. Secondly, the resolution simulated is for a grating where the lateral widths of the absorbers are thinner than the skin depth of the absorber metal. The resultant shallow depth of field at these grating periods ($\sim$ 3 nm), is admittedly restrictive, however it suggests that lithographers are only limited by processing constraints. As considerable as these constraints are, they are a challenge and not an impossibility.

In addition to investigations into what might fundamentally limit ENFOL's resolution, useful results that characterise evanescent exposures were obtained. The exposure of sub-wavelength gratings has been shown to be predominately due to TM polarised light. An enhanced, high contrast image of the grating is present in the near-field of the grating, due solely to the presence of evanescent diffracted orders. The depth of field is a critical parameter for an evanescent exposure. The reduction of depth of field with feature size is a common feature of all optical lithography systems, be it projection or contact lithography. The long-standing expression for defining the resolution of contact lithography, originally derived from the Fresnel relations, has been $l_{min} \sim \sqrt{\lambda g}$, where λ is the wavelength and g the mask–substrate gap (including photoresist). Solving the contact resolution equation in terms of g and redefining it to be the depth of field indicates a relationship proportional to l_{min}^2. Simulation results for sub-wavelength gratings reveal a more conservative linear relationship between the depth of field and grating period p. Furthermore, as $p \ll \lambda$, the depth of field appears to be independent of the exposure wavelength.

The near-field exposure shows a strong dependence on parameters such as grating period, duty cycle, and the radius of curvature of grating conductors. Shifts in the zeroth-order transmission coefficient may act to shift the optimal exposure threshold of different grating patterns on a mask. Clearly feature dependent exposures are undesirable. Ways to reduce these dependencies include operation at higher contrast factors, implying thinner resists. Another solution is to apply mask compensation techniques, such as grey-scale attenuators that could equalise the far-field light transmission.

A comprehensive treatment of evanescent exposures of grating structures has been presented, which has proved useful for understanding important concepts related to evanescent exposures with these geometries. A large set of geometries such as finite apertures of different shapes, and isolated, finite-length lines and shapes is yet to be explored. The evaluation of a broader set of geometries is an important extension to the work presented here. This would require a three-dimensional (3-D) rigorous electromagnetic simulator, although it is not obvious which technique will be most suitable. A 3-D multiple multipole program has been available in the past, and currently a new version is in the process of being updated and improved. However, the prospect of placing expansions in a 3-D space to obtain suitable models is daunting. Rigorous coupled wave analysis is inherently limited to periodic structures, although 2-D gratings (in 3-D space) can be simulated. 2-D gratings could be used to approximate exposures of isolated features and lines of finite dimensions, if large periods were modelled with respect to the feature of interest. Using large periods would minimise the effect of neighbouring features, although this is a somewhat artificial solution. A more suitable technique may be finite difference time domain methods, for which 2-D near field results have been reported [107, 108].

A novel photolithographic interference technique utilising evanescent light has been proposed – evanescent interference lithography (EIL). In this work simulations using the multiple multipole program show the potential of period halving in the near field by exposure of a high conductivity metallic grating, operating near a grating resonance with the grating period approximately equal to the effective wavelength in the surrounding medium.

Experimental demonstration of EIL is the next logical step from the EIL simulation work. An ideal experimental setup would consist of a polarised source, alignable to the mask, to filter out the TE polarisation that does not exhibit period halving. A collimated source with high spatial coherence is essential for an effect that relies on interference. A tunable narrow-band source is ideal so that the optimal resonant wavelength might be found for the physically fabricated grating. The material properties, surface quality and other parameters such as duty cycle of the grating can not generally be quantified exactly, leaving some doubt as to the size and position of the resonance behaviour predicted from simulations. Some control over the angle of the incident light may also be useful, although to replicate the EIL simulations only normally incident

light is required.

A further interesting extension of the evanescent near field work reported here would be an investigation into the area of superlenses reported by Pendry [155]. Superlenses consisting of a negative refractive index material are capable of perfect imaging. Pendry proposes that a slab of highly conducting metal with a negative permittivity will look like a superlens at optical wavelengths. This work has exciting implications for evanescent optical lithography, as analysis shows that the superlens amplifies the evanescent fields, including these components in the image reconstruction. This suggests the potential to eliminate depth of field issues.

Near-field optics is an exciting and fast developing field. The evanescent near field has been shown to provide access to resolutions far smaller than are attainable with conventional lens-based optical systems. The proposed technique, evanescent near-field optical lithography is able to capitalise on this additional resolution in a low-cost way. Further research in this area is likely to result in better process methods and improved understanding of the fundamental electromagnetic interactions that allow high resolution in this regime.

Appendix A

SIMULATIONS OF CURVED PHASE SHIFTING MASKS

A.1 INTRODUCTION

The last twenty years has seen some important changes in optical mask technology. Whereas traditionally optical mask technology manipulated just the amplitude of the electric field in a photolithographic exposure, today the phase is also manipulated. A combined amplitude and phase mask, the alternating-aperture phase-shifting mask (Alt-PSM) is now used for commercial IC manufacture [3, 156]. The principle of Alt-PSM is illustrated in Fig. A.1. A conventional amplitude mask is modified by the addition of a phase shifter to alternate apertures. The thickness of the shifter material is such that light passing through it is shifted by half a wavelength (π) with respect to light passing through unshifted areas. The light diffracted under the shadowed regions from each neighbouring aperture then arrives out of phase, interfering destructively instead of constructively.

Alt-PSM has been popular for manufacture of circuits such as DRAM which have a highly repetitive structure as it is easy to determine where to place the phase shifters. Using resolution enhancement techniques such as this pushes the k_1 factor (see Equ. 1.2) to even lower values, allowing the lifetime of current generation stepper technology to be extended.

Purely phase shifting masks, known as binary or unattenuating phase shifting masks also have useful exposure properties. The principle of binary PSM is illustrated in Fig. A.2. A surface relief pattern is created in a transparent phase shifting material. The depth u_1 of the surface relief is selected to produce a π phase shift. This equates to thicknesses

$$u_1 = \frac{N\lambda}{2\Delta n} \tag{A.1}$$

where N is an odd integer multiple, λ the illumination wavelength and Δn the difference in the refractive indices of the phase mask material and the surrounding material (in this case air). The light at either side of the edges in the surface relief arrives π out of

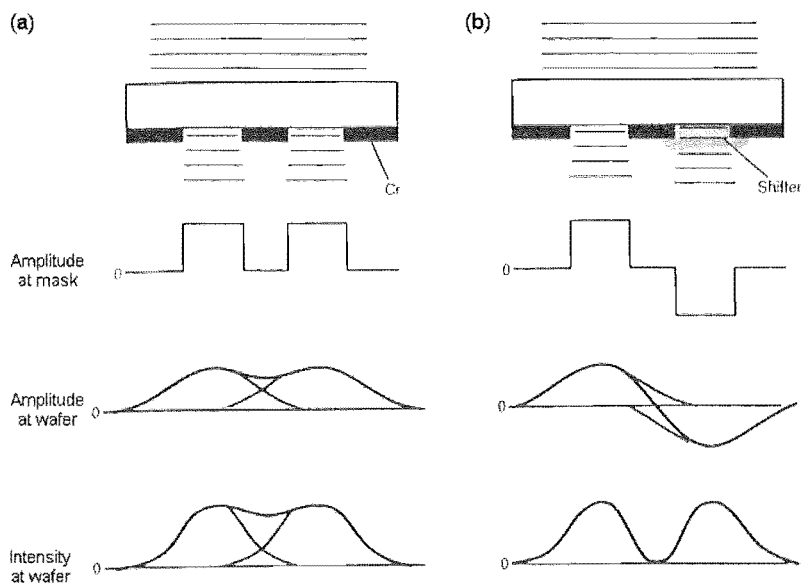


Figure A.1 Mask optics (a) Conventional amplitude mask. (b) Alternating-aperture phase-shifting mask. The shifter thickness is sufficient to introduce a half-wave of phase shift into light passing through it compared with light passing through an uncoated region. This has the effect of reversing the amplitude at one aperture compared to its neighbour converting constructive interference into destructive interference at the wafer. Reproduced from [156].

phase and interferes destructively producing the intensity profile as shown in the inset of Fig. A.2. The narrow intensity null can be captured in a photoresist film placed in the near field of the phase mask.

Sub-100 nm sized structures have been reported using this interference method with optical contact techniques [37], [39], and the technique has also been shown to be compatible with optical projection schemes [157]. The distinct advantage of this technique is that patterning of sub-100 nm scale structures can be obtained without the need for small features on the mask.

Previous reports of binary phase shifting masks have been for masks with steep sidewalls. In keeping with the low-cost, high resolution patterning theme, investigations of binary phase shifting masks with isotropic profiles typical of wet-etch processes were undertaken [39]. This appendix describes simulations which compare the differences between phase masks with profiles typical of anisotropic dry-etch techniques and that of the more accessible but isotropic profile of wet-etch techniques. This work supported experimental work that was able to produce sub-100 nm structures, using only a simple broad-band UV mask aligner, evaporator, and wet-bench.

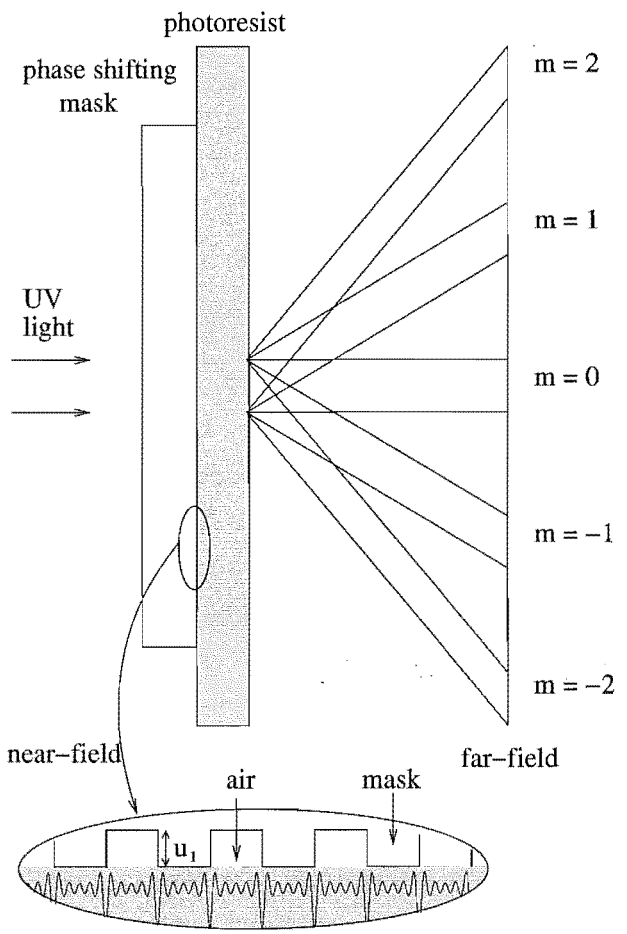


Figure A.2 Principle of binary PSM. Light is shone through a transparent but phase shifting material with a surface relief. In the mask's near field destructive interference occurs at the phase boundaries. The inset shows a close-up of the mask grating, and the corresponding intensity pattern. In the far-field a Fraunhofer diffraction pattern is evident.

A.2 SIMULATION METHOD

A simple scalar simulation method, previously used by Rogers *et al.* [37] was used to obtain the near-field and far-field intensity patterns for a variety of phase mask profiles. The procedure is summarised in Fig. A.3. Firstly an approximation of the electric field distribution at the exit plane of the mask is obtained. The electric field is assumed to be uniform, exhibiting only differences in phase across the mask, depending on the phase-shifter thickness. The electric field E is defined as

$$E = E_{in}e^{i\phi(x)} \tag{A.2}$$

where E_{in} is the incident electric field, and $\phi(x)$ is the phase. $\phi(x)$ is defined as

$$\phi(x) = \frac{2\pi}{\lambda}\Delta n u(x) \tag{A.3}$$

where Δn is the difference in refractive index between the mask material and the surrounding medium (typically air) and $u(x)$ defines the height of the surface relief as a function of x.

A Fourier transform of the approximated electric field distribution at the exit aperture of the PSM is computed to obtain the far-field diffraction pattern. The far-field distribution is then filtered to remove the high spatial frequencies, *i.e.*, $k > 2\pi/p$, where p is the period of the grating. These components are equivalent to the evanescent components. The inverse Fourier transform is then computed to obtain the near-field intensity profile, and the squared magnitude of this is used to obtain the relative intensity of the near-field illumination. Qualitatively this procedure is identical to the action of a convex optical lens which firstly collects the diffracted light, then focuses the image to reconstruct it.

The resultant near-field image equates to a Fresnel near-field image at the exit aperture, less any high-spatial-frequency evanescent components. As this method is scalar it is valid only where the period of the grating is greater than the illumination wavelength and when the mask-substrate gap is $> \lambda$, *i.e.*, outside the evanescent regime.

A.3 SIMULATION RESULTS

Simulations were performed to evaluate the transmitted intensity for phase mask gratings, for straight and curved sidewalls with periods from 1μm to 24μm, and etch depths from 0 to $2\,\mu$m. Silicon nitride was used as the mask material, with $n = 2$, surrounded by free space. The radius of curvature of the sidewalls was chosen to be equal to the maximum etch depth to approximate an isotropic wet-etch. Figure A.4 shows the near field intensity (solid line) for a $24\,\mu$m period curved phase mask profile with an etch depth of $0.59\,\mu$m. This etch depth was found to give the deepest intensity null at a

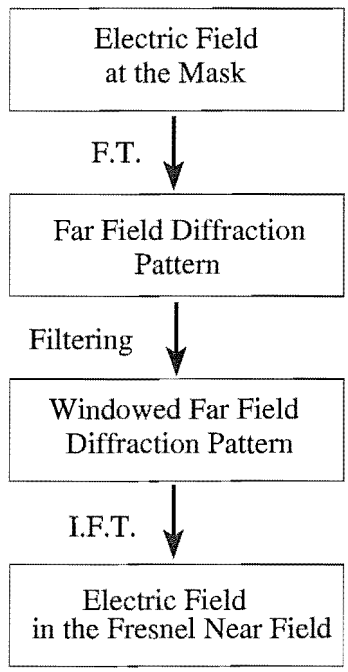


Figure A.3 Simulation method.

wavelength of 436 nm. The intensity for a square edged phase mask with the same etch depth is indicated by the dashed line. This depth does not produce a full π phase shift, hence the imperfect null. Note that the intensity profile for the curved phase mask is asymmetric, due to the mask's changing gradient. The minimum intensity is obtained at a position close to the steepest gradient of the mask. Evidence of this asymmetry has been observed as a shoulder in the photoresist profile for some exposures performed using masks etch with HF [39].

Results show that in order to obtain optimal contrast - ie a complete intensity null, a greater etch depth is required than for a straight-edged mask. Figure A.5(a) plots the minimum intensity obtained versus the etch depth for the curved profiles (solid line) and compares it with square edged phase masks (dashed line). The first complete intensity null for the square edged profile occurs at $0.22\,\mu$m and thereafter repeats itself at depths that coincide with a phase shift of an odd multiple of π. The first null for the curved profile does not occur until an etch depth of $0.59\,\mu$m and thereafter maintains a reasonably low intensity minimum. Near $2\,\mu$m the intensity minimum again comes very close to a complete null.

The shift of the intensity null to a deeper etch depth for the curved profile can be understood qualitatively by considering that the intensity null will occur at the steepest part of the curve. A complete null will not be formed until the radius is large enough to create a full π shift within a lateral distance small enough with respect

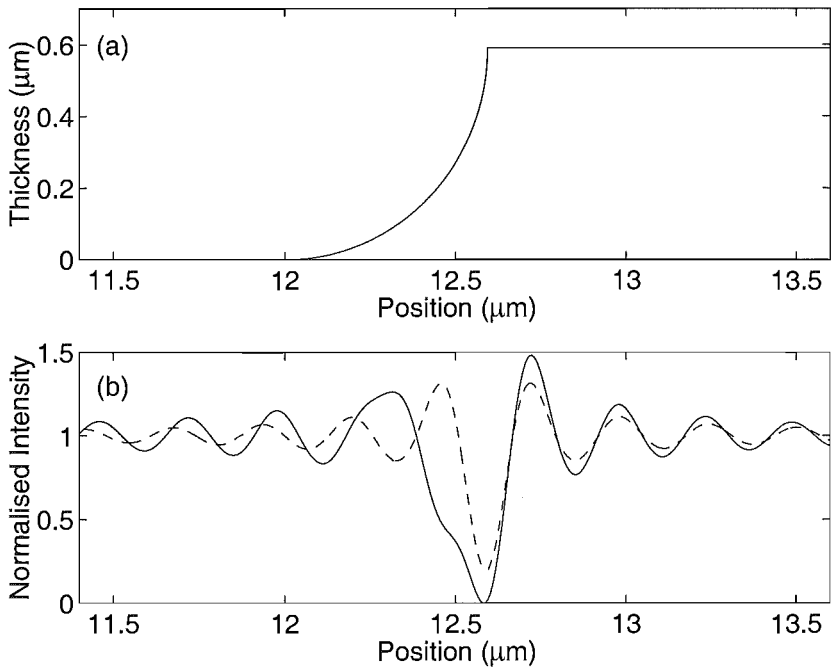


Figure A.4 Simulation result of the near-field intensity for a 436 nm exposure through a silicon nitride phase mask. (a) Profile of the curved phase mask, with 0.59 μm etch depth and radius. (b) Near-field intensity patterns for the curved phase mask (solid line) and a square-edged phase mask with the same etch depth (dashed line).

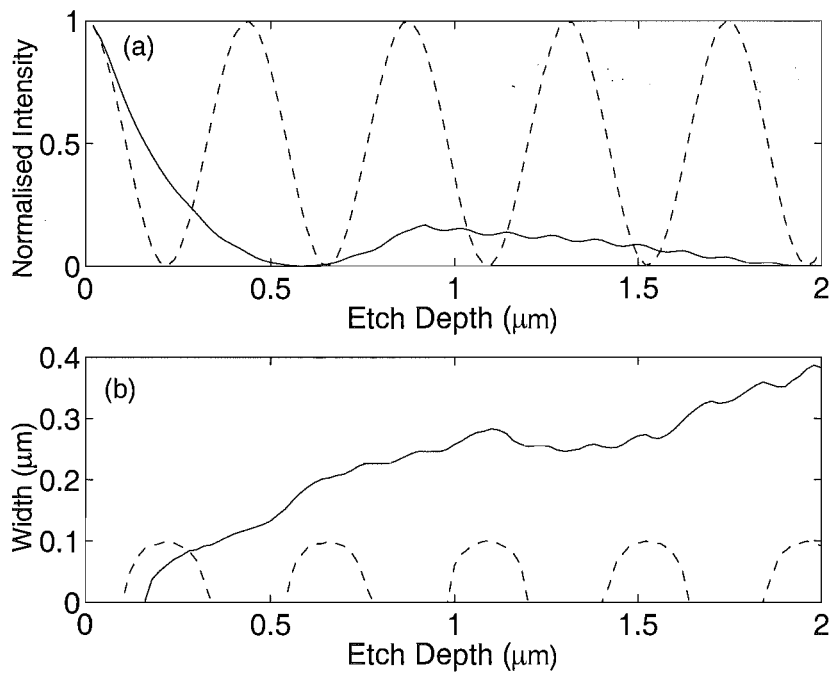


Figure A.5 Minimum intensity in the near-field, normalised to the intensity with no phase step for wet-etched phase masks with an isotropic etch profile (solid line) and square-edged phase masks (dashed line). (b) Full width at half maximum (FWHM) of the near field intensity minima for phase masks with wet-etched (solid line) and square-edged (dashed line) profiles.

to the illumination wavelength. Increasing the etch depth further does not totally destroy the null, as for the square-edged case, since there is always somewhere along the curved profile which offers a phase change close to π and hence results in destructive interference. A sloped profile of 45° for the same etch depth does not result in a significant intensity null as the gradient is nowhere large enough to create a phase change close to π.

Another interesting result is the potential to control the linewidth using a curved profile, as can be seen in Fig. A.5(b). This plots the full width at half maximum (FWHM) of the null of the near field intensity profile. With straight sidewall profiles (dashed line), the FWHM repeats itself periodically as the etch depth cycles through odd multiples of π phase shifts, with the maximum obtainable FWHM being approximately $\lambda/4$. With the curved profile (solid line) there is a steady broadening of the linewidth, as more areas along the curve obtain a steep enough gradient to create a significant phase shift. Curved phase mask profiles therefore offer the possibility of producing a range of feature sizes for a given combination of photoresist and illumination wavelength.

A.4 SUMMARY

Simulations of phase masks with curved profiles, compatible with wet-etching techniques show that nano-scale lithography is possible over a broad-range of etch depths. With curved profiles, a deeper etch depth is required, compared to the straight-edged PSM to ensure that a steep enough phase gradient is obtained over a enough thickness to create a deep enough intensity null. A broader intensity null is predicted with the curved phase masks, as well as some degree of linewidth control via the etch depth.

REFERENCES

[1] H. Queisser. *The Conquest of the Microchip.* Harvard University Press, 1988.

[2] E Korcynski. Moore's law extended: The return of cleverness. *Solid State Technology*, pages 359–364, July 1997. Interview with Gordon Moore.

[3] The international technology roadmap for semiconductors. http://www.sematech.org, 1999.

[4] H. I. Smith. *Submicron-and Nanometer-Structures Technology.* Massachusetts Institute of Technology, Cambridge, MA 02139. Lecture notes for Course 6.781, Department of Electrical Engineering and Computer Science.

[5] P. Packan. Fundamental issues of scaling CMOS devices. Stanford Instructional Television Network Stanford Center for Professional Development, 1999/2000. EE380 Seminar, Standford University.

[6] G. L.-T. Chiu and J. M. Shaw. Optical lithography: Introduction. *IBM J. Res. Develop.*, 41:3–6, 1997.

[7] H. H. Solak, D. He, W. Li, and F. Cerrina. Nanolithography using extreme ultraviolet lithography interferometry: 19 nm lines and spaces. *J. Vac. Sci. Technol. B*, 17(6):3052–3057, 1999.

[8] S. T. Stanton, J. A. Liddle, W. K. Waskiewicz, and M. M. Mkrtchyan. Critical tool performance analysis for SCALPEL extensibility. *J. Vac. Sci. Technol. B*, 18(1):112–116, 2000.

[9] H. C. Pfeiffer, R. S. Dhaliwal, S. D. Golladay, S. K. Doran, M. S. Gordon, T. R. Groves, R. A. Kendall, J. E. Lieberman, P. F. Petric, D. J. Pinckny, R. J. Quickle, C. F. Robinson, J. D. Rockrohr, J. J. Senesi, W. Stickel, E. V. Tressler, A. Tanimoto, T. Yamaguchi, K. Okamoto, K. Suzuki, T. Okino, S. Kawata, K. Morita, S. C. Suziki, H. Shimizu, S. Kojima, G. Varnell, W. T. Novak, D. P. Stumbo, and M. Sogard. Projection reduction exposure with variable axis immersion lenses: Next generation lithography. *J. Vac. Sci. Technol. B*, 17(6):2840–2856, 1999.

[10] R. H. Austin. Biomolecules and microdevices. In *The 44th International Conference on Electron, Ion and Photon Beam Technology and Nanofabrication*, page 3, Palm Springs, CA, USA, 2000.

[11] H. G. Craighead, B. Ilic, D. Czaplewski, and P. Neuzil. Resonant biosensor for the detection of escherichia coli. In *The 44th International Conference on Electron, Ion and Photon Beam Technology and Nanofabrication*, pages 393–394, Palm Springs, CA, USA, 2000.

[12] S. M. Sze. *Semiconductor Devices Physics and Technology*, chapter 11. John Wiley & Sons, New York, 1985.

[13] C. Wagner, W. Kaiser, J. Mulkens, and D. G. Flagello. The technical considerations of extending optical lithography. *Solid State Technology*, pages 97–108, September 2000.

[14] R. W. Hardin. Lasers in lithography: A race against time. *Phot. Spectra*, pages 94–100, October 1997.

[15] H. I. Smith and M. L. Schattenburg. X-ray lithography from 500 to 30 nm: X-ray nanolithography. *IBM J. Res. Develop.*, 37(3):319–329, 1993.

[16] H. I. Smith. X-ray lithography for microelectronics. *Physica Scripta.*, T61:26–31, 1996.

[17] L. Harriott and A. Liddle. Electron-beam lithography. *Physics World*, pages 41–45, April 1997.

[18] W. M. Moreau. *Semiconductor lithography: Principles, Practices & Materials*, chapter 9. Plenum Press, New York, 1988.

[19] H. Pfeiffer. Variable spot shaping for electron beam lithography. *J. Vac. Sci. Technol.*, 15:887, 1978.

[20] T. H. P. Chang, M. G. R. Thomson, E. Kratchmer, Kim H. S., M. L. Yu, and K. Y. Lee. Electron-beam microcolumns for lithography and related applications. *J. Vac. Sci. Technol. B*, 14(6):3774–3781, 1996.

[21] L. P. Muray, K. Y. Lee, J. P. Spallas, M. Mankos, Y. Hsu, M. R. Gmur, H. S. Gross, C. B. Stebler, and T. H. P. Chang. Experimental evaluation of arrayed microcolumn lithography. *Microelectronic Engineering*, 53:271–277, 2000.

[22] L. Harriott, W. Waskiewicz, A. Novembre, and J. A. Liddle. Favored SCALPEL's continued progress. *Solid State Technology*, page 73, July 1999.

[23] L. Fetter, C. Biddick, M. Blakey, A. Liddle, M. Peabody, A. Novembre, and D. Tennant. Patterning of membrane masks for projection e-beam lithography. *Proc. SPIE*, 2884, 1996.

[24] J. E. Bjorkholm. EUV lithography - the successor to optical lithography. *Intel Technology Journal*, 1998.

[25] S. Y. Chou, P. R. Krauss, and P. J. Renstrom. Imprint of sub-25 nm vias and trenches in polymers. *Appl. Phys. Lett.*, 67(21):3114–16, November 1995.

[26] S. Y. Chou. Patterned magnetic nanostructures and quantized magnetic disks. *Proceedings of the IEEE*, 85(4):652–671, April 1997.

[27] S. Y. Chou, P. R. Krauss, W. Zhang, L. Guo, and L. Zhuang. Sub-10 nm imprint lithography and applications. *J. Vac. Sci. Technol. B*, 15(6):2897–2904, 1997.

[28] S. Y. Chou, P. R. Krauss, and P. J. Renstrom. Nanoimprint lithography. *J. Vac. Sci. Technol. B*, 14(6):4129–4133, 1996.

[29] J. Haisma, M. Verheijen, K. van den Heuvel, and J. van den Berg. Mold-assisted nanolithography: A process for reliable pattern replication. *J. Vac. Sci. Technol. B*, 14(6):4124–4128, 1996.

[30] M. M. Alkaisi, R. J. Blaikie, and S. J. McNab. Low temperature nanoimprint lithography using silicon nitride molds. In *Micro- and Nano-Engineering 2000*, Jena, Germany, 18-21 September, 2000.

[31] D. Khang and H. H. Lee. Room-temperature imprint lithography by solvent vapor treatment. *Appl. Phys. Lett.*, 76(7):870–872, February 2000.

[32] A. Kumar and G. Whitesides. Features of gold having micrometer to centimeter dimensions can be formed through a combination of stamping with an elastomeric stamp and an alkanethiol "ink" followed by chemical etching. *Appl. Phys. Lett.*, 63(14):2002–2004, October 1993.

[33] H. A. Biebuyck, N. B. Larsen, E. Delamarche, and B. Michel. Lithography beyond light: Microcontact printing with monolayer resists. *IBM J. Res. Develop.*, 41(1):159–170, 1997.

[34] Y. Chen, A. Lebib, F. Carcenac, and H. Launois. Microcontact printing and pattern transfer with a tri-layer processing. *Microelectronic Engineering*, 53:253–256, 2000.

[35] Y. Xia, E. Kim, J. Zhao, X. Rogers, M. Prentiss, and G. Whitesides. Complex optical surfaces formed by replica molding against elastomeric masters. *Science*, 273:347–349, July 1996.

[36] D. Wang, S. Thomas, K. Wang, and Y. Xia. Nanometer scale patterning & pattern transfer on amorphous Si, crystalline Si, & SiO_2 surfaces using self-assembled monolayers. *Appl. Phys. Lett.*, 70(12):1593–1595, March 1997.

[37] J. A. Rogers, K. E. Paul, R. J. Jackman, and G. M. Whitesides. Using an elastomeric phase mask for sub-100nm photolithography in the optical near-field. *Appl. Phys. Lett.*, 70(20):2658–2660, May 1997.

[38] J. A. Rogers, K. E. Paul, and G. M. Whitesides. Quantifying distortions in soft lithography. *J. Vac. Sci. Technol. B*, 16(1):88–97, 1998.

[39] M. M. Alkaisi, R. J. Blaikie, and S. J. McNab. Nanolithography using wet etched silicon nitride phase masks. *J. Vac. Sci. Technol. B*, 16(6):3929–3933, 1998.

[40] H. Chen, R. J. Blaikie, R. Cheung, and M. M. Alkaisi. Nanolithography using reactive ion etched silicon nitride phase masks. In *Proceedings of the Sixth Electronics New Zealand Conference*, pages 125–130, Auckland, New Zealand, September 1999.

[41] M. D. Levenson, N. S. Viswanathan, and R. A. Simpson. Improving resolution in photolithography with a phase-shifting mask. *IEEE Transactions of Electron Devices*, 29(12):1812–1846, December 1982.

[42] H. I. Smith. Method for fabricating high frequency surface wave transducers. *Rev. Sci. Instrum.*, 40:729–730, 1969.

[43] H. I. Smith, N. Efremow, and P. L. Kelly. Photolithographic contact printing of 4000Å linewidth patterns. *J. Electrochemical Soc.*, 121(11):1503–1506, 1974.

[44] H. I. Smith. Fabrication techniques for surface-accoustic-wave and thin-film optical devices. *Proceedings of the IEEE*, 62(10):1361–1387, October 1974.

[45] T. Ono and M Esashi. Subwavelength pattern transfer by near-field photolithography. *Japanese J. of Applied Physics*, 37(12B):6745–6749, December 1998.

[46] J. G. Goodberlet and L. D. Bryan. Deep-ultraviolet contact photolithography. *Microelectronic Engineering*, 53:95–99, 2000.

[47] J. G. Goodberlet. Patterning 100 nm features using deep-ultraviolet contact photolithography. *Appl. Phys. Lett.*, 76(6):667–669, 2000.

[48] J. G. Goodberlet. Private communication.

[49] H. Schmid, H. Biebuyck, B. Michel, and O. J. F. Martin. Light-coupling masks for lensless, sub-wavelength optical lithography. *Appl. Phys. Lett.*, 72(19):2379–2381, May 1998.

[50] H. Schmid, H. Biebuyck, B. Michel, O. J. F. Martin, and N. B. Piller. Light-coupling masks: An alternative, lenseless approach to high-resolution optical contact lithography. *J. Vac. Sci. Technol. B*, 16(6):3422 – 3425, 1998.

[51] R.J. Blaikie, M.M. Alkaisi, S.J. McNab, D.R.S. Cumming, and D.G. Hasko. Nanolithography using optical contact exposure in the evanescent near field. *Microelectronic Engineering*, 46:85–88, 1999.

[52] M. M. Alkaisi, R. J. Blaikie, S. J. McNab, R. Cheung, and D.R.S Cumming. Sub-diffraction-limited patterning using evanescent near-field optical lithography. *Appl. Phys. Lett.*, 75(22):3560–3562, 1999.

[53] M. M. Alkaisi, R. J. Blaikie, and S. J. McNab. 70 nm features on 140 nm period using evanescent near field optical lithography. *Microelectronic Engineering*, 53:237–240, 2000.

[54] S. J. McNab and R. J. Blaikie. Contrast in the evanescent near field of $\lambda/20$ period gratings for photolithography. *Applied Optics*, 39:20 – 25, January 2000.

[55] G. O. Reynolds, J. B. DeVelis, G. B. Parrent, and B. J. Thompson. *The New Physical Optics Notebook: Tutorials in Fourier Optics.* SPIE - The International Society for Optical Engineering, 1989.

[56] L. F. Thompson, C. G. Wilson, and M. J. Bowden, editors. *Introduction to Microlithography.* American Chemical Society, Washington, DC, second edition, 1994.

[57] E. Betzig, J. K. Trautman, T. D. Harris, J. S. Weiner, and R. L. Kostelak. Breaking the diffraction barrier: Optical microscopy on a nanometric scale. *Science*, 251:1468–1470, 1991.

[58] Motoichi Ohtsu and Hirokazu Hori. *Near-field nano-optics, From basic principles to nano-fabrication and nano-photonics.* Kluwer Academic/Plenum Publishers, 1999.

[59] E. Betzig, P. L. Finn, and J. S. Weiner. Combined shear force and near-field scanning optical microscopy. *Appl. Phys. Lett.*, 60(20):2484–2486, May 1992.

[60] I. I. Smolyaniinov, D. L. Mazzoni, and C. C. Davis. Near-field direct-write ultraviolet lithography and shear force microscopic studies of the lithographic process. *Appl. Phys. Lett.*, 67(25):3859–3861, December 1995.

[61] S. Madseen, M. Mullenborn, K. Birkelund, and F. Grey. Optical near-field lithography on hydrogen-passivated silicon surfaces. *Appl. Phys. Lett.*, 69(4):544–546, July 1996.

[62] M. Tabe and T. Yamamoto. Nanometer-scale local oxidation of silicon using silicon nitride islands formed in the early stages of nitridation. *Appl. Phys. Lett.*, 69(15):2222–2224, 1996.

[63] K. Wilder and C. F Quate. Scanning probe lithography using a cantilever with integrated transistor for on-chip control of the exposing current. *J. Vac. Sci. Technol. B*, 17(6):3256–3261, 1999.

[64] R. Lane. Optical engineering. Course Notes for 641, Department of Electrical & Electronic Engineering, University of Canterbury, 1999.

[65] T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff. Extraordinary optical transmission through sub-wavelength hole arrays. *Nature*, 391:667–669, February 1998.

[66] H. A. Bethe. Theory of diffraction by small holes. *Physical Review*, 66:163–182, October 1944.

[67] M. M. J. Treacy. Dynamical diffraction in metallic optical gratings. *Appl. Phys. Lett.*, 75(5):606–608, August 1999.

[68] E. Loewen and E. Popov. *Diffraction Gratings and Applications.* Marcel Dekker, New York, 1997. Chapter 2.

[69] L. Reimer. *Scanning Electron Microscopy: Physics of Image Formation and Microanalysis*, chapter 1, 5. Springer, 1998.

[70] Philips. *Operating Instructions PSEM500*, first edition.

[71] D. Bollinger, S. Iida, and O. Matsumoto. Reactive ion etching: Its basis and future, Part I. *Solid State Technology*, 27:111–117, May 1984.

[72] Dennis M. Manos and Daniel L. Flamm, editors. *Plasma Etching An Introduction.* Plasma - Materials Interactions. Academic Press Inc., 1989.

[73] Scanning probe microscopy training notebook. Digital Instruments.

[74] R. Howland and L. Benatar. A practical guide to scanning probe microscopy. Park Scientific Instruments.

[75] R. Stevens, C. Nguyen, A. Cassell, L. Delzeit, M. Meyyappan, and J. Han. Improved fabrication approach for carbon nanotube devices. *Appl. Phys. Lett.*, 77(21):3453–3455, November 2000.

[76] L. E. Trimble, G. K. Celler, J. Frackoviak, and G. R. Weber. Membrane fragility: Fact or illusion? *J. Vac. Sci. Technol. B*, 10(6):3200–3203, 1992.

[77] W. M. Moreau. *Semiconductor lithography: Principles, Practices & Materials.* Plenum Press, New York, 1988.

[78] M. J. Rooks, C. C. Eugster, J. A. del Alamo, G. L. Snider, and E. L. Hu. Split-gate electron waveguide fabrication using multilayer poly(methylmethacrylate). *J. Vac. Sci. Technol. B*, 9(6):2856–2860, 1991.

[79] C. Raetzel, S. Schild, and H. Schlotterer. Etching properties of different silicon films. In *Meeting of the Electrochemical Society*, pages 136–138, 1974. Extended abstract.

[80] M. J. Rooks and R. C. Tiberio. Coherence of large gratings and electron-beam fabrication techniques for atom-wave interferometry. *J. Vac. Sci. Technol. B*, 13(6):2745–2751, 1995.

[81] G. Kaminsky. Micromachining of silicon mechanical structures. *J. Vac. Sci. Technol. B*, 3(4):1015–1024, 1985.

[82] J. Aizenberg, J.A. Rogers, and K.E. Paul. Imaging the irradiance distribution in the optical near field. *Appl. Phys. Lett.*, 71(26):3773–3775, December 1997.

[83] S. J. McNab and R. J. Blaikie. Investigating the fundamental limit to resolution in the evanescent near field. In *Proceedings of Image & Vision Computing New Zealand*, pages 97 – 102, Christchurch, NZ, 30-31 August, 1999.

[84] H. Devereux and R. Cheung. Low temperature reactive ion etching of silicon for micromachining. In *Proceedings of the Fifth Electronics New Zealand Conference*, Dunedin, New Zealand, 31 August - 1 September, 1998.

[85] W. M. Moreau. *Semiconductor lithography: Principles, Practices & Materials*, page 567. Plenum Press, New York, 1988.

[86] S. P. Lyman, J. L. Jackel, and P. L. Liu. Lift-off of thick metal layers using multilayer resist. *J. Vac. Sci. Technol.*, 19(4):1325–1328, 1981.

[87] A. A. Milgram. Lift-off process for achieving fine-line metallization. *J. Vac. Sci. Technol. B*, 1(2):490–493, 1983.

[88] B. J. Lin, V. W. Chao, K. E. Petrillo, and B. J. L. Yang. A molded deep uv portable conformable masking system. *Polymer Engineering & Science*, 1986.

[89] J. R. Lothian, S. J. Ren, F. Pearton, U. K. Chakrabarti, C. R. Abernathy, and A. Katz. Trilayer lift-off metallization process using low temperature deposited SiN_x. *J. Vac. Sci. Technol. B*, 10(6):2361–2365, 1992.

[90] M. L. Schattenburg, R. J. Aucoin, and R. C. Fleming. Optically matched trilevel resist process for nanostructure fabrication. *J. Vac. Sci. Technol. B*, 13(6):3007–3011, 1995.

[91] V. Bouchiat and D. Esteve. Lift-off lithography using an atomic force microscope. *Appl. Phys. Lett.*, 69(2):398–100, November 1996.

[92] R. W. Ryan, R. F. Kopf, R. A. Hamm, R. J. Malik, R. Masaitis, and R. Opila. Dielectric-assisted trilayer lift-off process for improved metal definition. *J. Vac. Sci. Technol. B*, 16(5):2759–2762, 1998.

[93] H. Jansen, Meint de Boer, and *et al.*. The black silicon method II: The effect of mask material and loading on the reactive ion etching of deep silicon trenches. *Microelectronic Engineering*, 27:475–480, 1995.

[94] David R. Lide, editor. *Handbook of Chemistry and Physics.* CRC Press, Boca Raton, 75th edition, 1994.

[95] S. Ramo, J.R. Whinnery, and T. Van Duzer. *Fields and Waves in Communication Electronics*, page 294. Wiley, second edition, 1984.

[96] K. E. Petrillo, V. W. Chao, and B. J. Lin. Submicrometer contact hole delineation with a two-layer deep-UV portable conformable masking system. *J. Vac. Sci. Technol. B*, 1(4):1219–1224, 1983.

[97] R. P. Baumann. *Absorption Spectroscopy.* John Wiley & Sons Inc., 1962.

[98] Ulrich Brackmann. Blue green dyes. Published on web for Lambdachrome Laser Dyes.

[99] Brewer Science Inc. ARC data sheet for XL family.

[100] Gsolver 4.20a. Grating Solver Development Company. Tables of refractive indices from Sopra (http://www.gsolver.com).

[101] H. Namatsu, M. Nagase, T. Yamaguchi, K. Yamazaki, and K. Kurihara. Influence of edge roughness in resist patterns on etched patterns. *J. Vac. Sci. Technol. B*, 16(6):3315–3321, 1998.

[102] Shipley representative Gerhard Bleidießel. Private communication.

[103] M. Born and E. Wolf. *Principles of Optics.* Pergamon Press, sixth edition, 1980.

[104] B. J. Lin. Electromagnetic near-field diffraction in a medium slit. *J. of the Optical Soc. of America*, 62(8):976–981, August 1972.

[105] D. W. Prather. Techniques for applying the FDTD method to the electromagnetic analysis of diffractive optical elements. In *Diffractive Optics and Micro-Optics*, pages 5–9, Quebec City, Canada, 18-22 June, 2000. Optical Society of America.

[106] A. Chavez-Pirson and S. T. Chu. A full vector analysis of near-field luminescence probing of a single quantum dot. *Appl. Phys. Lett.*, 74(11):1507–1509, 1999.

[107] H. Ichikawa and H. Kikuta. Electromagnetic diffraction and subwavelength mask lithography: numerical feasibility study. In *Diffractive Optics and Micro-Optics*, pages 16–18, Quebec City, Canada, 18-22 June, 2000. Optical Society of America.

[108] S. Tanaka, M. Nakao, M. Umeda, K. Ito, S. Nakamura, and Y. Hatamura. Simulation of near-field photolithography using finite-difference time-domain method. To be published in Appl. Phys. Lett.

[109] T. H. Hubing. Survey of numerical electromagnetic modeling techniques. Technical report, University of Missouri-Rolla, 1991.

[110] Hewlett-Packard Company, 3000 Hanover Street, Palo Alto, CA 94304 USA. *High-frequency structure simulator User's Reference*, 1997.

[111] M. G. Moharam, Eric B. Gran, Drew A. Pommet, and T. K. Gaylord. Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings. *J. Opt. Soc. Am. A*, 12(5):1068–1076, May 1995.

[112] L. Li and C. W. Haggans. Convergence of the coupled-wave method for metallic lamellar diffraction gratings. *J. of the Optical Soc. of America A*, 10:1184–1189, 1993.

[113] P. Lalanne and G. M. Morris. Highly improved convergence of the coupled-wave method for TM polarization. *J. of the Optical Soc. of America A*, 13(4):779–784, April 1996.

[114] Gsolver. Grating Solver Development Company. (http://www.gsolver.com).

[115] O. J. F. Martin, C. Girard, and A. Dereux. Generalized field propagator for electromagnetic scattering and light confinement. *Physical Review E*, 74(4):525–529, January 1995.

[116] O. J. F. Martin and N. B. Piller. Electromagnetic scattering in polarizable backgrounds. *Physical Review E*, 58(3):3909–3915, September 1998.

[117] H. Singer, H. Steinbigler, and P. Weiss. A charge simulation method for the calculation of high-voltage fields. *IEEE Trans. on Power Appar. Syst.*, 93:1660–1668, 1974.

[118] C. Hafner. Multiple multipole (MMP) computations of guided waves and waveguide discontinuities. *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, 3:247–57, 1990. Chichester, J. Wiley.

[119] C. Hafner. *The Generalized Multipole Technique for Computational Electromagnetics.* Artech House Inc., Boston, 1990.

[120] C. Hafner. Multiple multipole program computation of periodic structures. *J. of the Optical Soc. of America A*, 12(5):1057–1067, May 1995.

[121] C. H. Hafner. *MaX-1 A Visual Electromagnetics Platform.* Wiley, Chichester, UK, 1998.

[122] Ashcroft and Merman, editors. *Properties of Silicon.* INSPEC IEE, London, 1988.

[123] G. Tayeb. The method of fictitious sources applied to diffraction gratings. *Appl. Comput. Electromag. Soc.*, 9(3):90–100, 1994.

[124] D. C. Bao, G. Dobson and J. A. Cox. Mathematical studies in rigorous grating theory. *J. Opt. Soc. Am. A*, 12(5):1029–1042, May 1995.

[125] C. H. Hafner. *Post-Modern Electromagnetics Using Intelligent Maxwell Solvers.* Wiley, Chichester, 1999.

[126] E. Loewen and E. Popov. *Diffraction Gratings and Applications.* Marcel Dekker, New York, 1997.

[127] M. K. Herndon, R. T. Collins, R. E. Hollingsworth, P. R. Larson, and M. B. Johnson. Near-field scanning optical nanolithography using amorphous silicon photoresists. *Appl. Phys. Lett.*, 74(1):141–143, January 1999.

[128] S. Friebel, J. Aizenberg, S. Abad, and P. Wiltzius. Ultraveiolet lithography of self-assembled monolayers for submicron patterned deposition. *Appl. Phys. Lett.*, 77(15):2406–2408, October 2000.

[129] S. Ramo, J.R. Whinnery, and T. Van Duzer. *Fields and Waves in Communication Electronics*, page 65. Wiley, second edition, 1984.

[130] S. Ramo, J.R. Whinnery, and T. Van Duzer. *Fields and Waves in Communication Electronics.* Wiley, second edition, 1984.

[131] Institute of Electrical Engineers Information Services Division. *Properties of Silicon.* London; INSPEC, Institution of Electrical Engineers, 1988.

[132] H. F. Talbot. Facts relating to optical science. *Phil. Mag.*, 9:401–407, 1836.

[133] Lord Rayleigh. On copying diffraction-gratings, and on some phenomena connected therewith. *Phil. Mag.*, 11(67), 1880.

[134] D. E. Grupp, H. J. Lezec, T. W. Ebbesen, K. M. Pellerin, and T. Thio. Crucial role of metal surface in enhanced transmission through subwavelength apertures. *Appl. Phys. Lett.*, 77(11):1569–1571, 2000.

[135] R. J. Blaikie and S. J. McNab. Evanescent interferometric lithography. *Applied Optics*, 40(10):1692–1698, April 2001.

[136] E. Hecht. *Optics*, pages 108–113. Addison-Wesley, second edition, 1987.

[137] D. C. Flanders, A. M. Hawryluk, and H. I. Smith. Spatial period division - a new technique for exposing submicrometer-linewidth periodic and quasiperiodic patterns. *J. Vac. Sci. Technol.*, 16(6):1949–1952, 1979.

[138] M. K. Berry and S. Klein. Integer, fractional and fractal Talbot effects. *J. of Modern Optics*, 43(10):2139–2164, 1996.

[139] E. Yablonovitch, T. J. Gmitter, and K. M. Leung. Photonic band gap: The face-centered-cubic case employing nonspherical atoms. *Phys. Rev. Lett.*, 67(17):2295–2298, 1991.

[140] C. Cuisin, Y. Chen, D. Decanini, A. Chelnokov, F. Carcenac, A. Madouri, J. Lourtioz, and H. Launois. Fabrication of three-dimensional microstructures by high resolution x-ray lithography. *J. Vac. Sci. Technol. B*, 17(6):3444–3448, 1999.

[141] D. M. Tennant, T. L. Koch, P. P. Mulgrew, and R. P. Gnall. Characterization of near-field holography grating masks for optoelectronics fabricated by electron beam lithography. *J. Vac. Sci. Technol. B*, 10(6):2530–2535, 1992.

[142] D. M. Tennant, K. F. Dreyer, K. Feder, and *et al.*. Advances in near field holographic grating mask technology. *J. Vac. Sci. Technol. B*, 12(6):3689–3694, 1994.

[143] K. O. Hill, B. Malo, F. Bilodeau, D. C. Johnson, and J. Albert. Bragg gratings fabricated in monomode photosensitive optical fiber by UV exposure through a phase mask. *Appl. Phys. Lett.*, 62(10):1035–1037, March 1993.

[144] T. J. Suleski, B. Baggett, W. F. Delaney, C. Koehler, and E. G. Johnson. Fabrication of high-spatial-frequency gratings through computer-generated near-field holography. *Optics Letters*, 24(9):602–604, May 1999.

[145] Karl Suss. Near field holography. http://www.suss.com/sitemap/near_field.htm, May 2000.

[146] S. H. Zaidi and R. J. Brueck. Multiple-exposure interferometric lithography. *J. Vac. Sci. Technol. B*, 11(3):658–666, 1993.

[147] S. H. Zaidi and R. J. Brueck. Multiple exposure interferometric lithography. *Proc. SPIE*, 2197:869–875, 1994.

[148] E. Loewen and E. Popov. *Diffraction Gratings and Applications.* Marcel Dekker, New York, 1997. Chapter 8.

[149] H. Raether. *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*, volume 111 of *Springer Tracts in Modern Physics.* Springer-Verlag, 1988.

[150] Clariant. Resist properties. http://www.azresist.com, 2000.

[151] B. Stenkamp, M. Abraham, W. Ehrfeld, E. Knapek, M. Hintermaier, M. Gale, and R. Morf. Grid polarizer for the visible spectral region. *Proc. SPIE*, 2213:288–296, 1994.

[152] Z. Yu, P. Deshpande, W. Wu, J. Wang, and S. Chou. Reflective polarizer based on a stacked double-layer subwavelength metal grating structure fabricated using nanoimprint lithography. *Appl. Phys. Lett.*, 77(7):927–929, August 2000.

[153] T. Doumuki and H. Tamada. An aluminium-wire grid polarizer fabricated on a gallium-arsenide photodiode. *Appl. Phys. Lett.*, 71(5):686–688, 1997.

[154] M. B. Stern, S. C. Palmateer, M. W. Horn, M. Rothschild, B. E. Maxwell, and J. E. Curtin. Profile control in dry development of high-aspect-ratio resist structures. *J. Vac. Sci. Technol. B*, 13(6):3017–3021, 1995.

[155] J. B. Pendry. Negative refraction makes a perfect lens. *Phys. Rev. Lett.*, 85(18):3966–3969, October 2000.

[156] M. D. Levenson. Wavefront engineering for photolithography. *Physics Today*, pages 29–36, 1993.

[157] H. Watanabe, Y. Todokoro, and M. Inoue. Transparent phase shifting mask. *IEEE IEDM*, 90:821–824, 1990.